CU12-H03045-17-P00024?
657.05
AC27

THE ACCOUNTING REVIEW

Business Studies

17

VOLUME 67, No. 1 JANUARY 1992

HERBERT F. TAGGART, NORTON M. BEDFORD, R. J. CHAMBERS, SIDNEY DAVIDSON, CARL THOMAS DEVINE, THOMAS R. DYCKMAN, YUJI IJIRI, THOMAS F. KELLER, and ROBERT G. NICHOLS
A Tribute to William A. Paton
DAVID A. GUENTHER
Taxes and Organizational Form: A Comparison of Corporations and Master Limited Partnerships
DAVID N. RICCHIUTE
Working-Paper Order Effects and Auditors' Going-Concern Decisions
LISA KOONCE
Explanation and Counterexplanation During Audit Analytical Review
STEVEN F. CAHAN
The Effect of Antitrust Investigations on Discretionary Accruals: A Refined Test of the Political-Cost Hypothesis

A FORUM ON LABORATORY MARKETS AND AUDITING RESEARCH
NICHOLAS DOPUCH and RONALD R. KING
Negligence Versus Strict Liability Regimes in Auditing: An Experimental Investigation
DAVID E. WALLIN
Legal Recourse and the Demand for Auditing
THE ARIZONA STATE UNIVERSITY RESEARCH WORKSHOP
JAMES R. BOATSMAN, LAWRENCE P. GRASSO, MICHAEL B. ORMISTON, J. HAL RENEAU
A Perspective on the Use of Laboratory Market Experimentation in Auditing Research
THE UNIVERSITY OF IOWA ACCOUNTING WORKSHOP
DOUGLAS V. DEJONG and ROBERT FORSYTHE
A Perspective on the Use of Laboratory Market Experimentation in Auditing Research

SHORTER ARTICLES
JENNIFER J. GAVER, KENNETH M. GAVER, and GEORGE P. BATTISTEL
The Stock Market Reaction to Performance Plan Adoptions
ASHIQ ALI, APRIL KLEIN, and JAMES ROSENFELD
Analysts' Use of Information about Permanent and Transitory Earnings Components in Forecasting Annual EPS
MICHAEL W. MAHER, PETER TIESSEN, ROBERT COLSON, and AMY J. BROMAN
Competition and Audit Fees

A QUARTERLY JOURNAL OF THE
AMERICAN ACCOUNTING ASSOCIATION

AMERICAN ACCOUNTING ASSOCIATION

The American Accounting Association was founded in 1916 as the American Association of University Instructors in Accounting. Its present name was adopted in 1936. The Association is a voluntary organization of persons interested in accounting education and research. Full membership entitles one to receive *The Accounting Review*, *Accounting Education News*, *Accounting Horizons*, and *Issues in Accounting Education*, and to purchase Association publications at a discount.

The members of the 1990–91 Executive Committee of the Association are as follows:

President	Arthur Wyatt, Arthur Andersen & Company
President-Elect	Gary L. Sundem, University of Washington
Vice President	Mark A. Wolfson, Stanford University
Vice President	David A. Wilson, Ernst & Young
Vice President	Andre Zund, University of St. Gall-Switzerland
Secretary-Treasurer	Jane O. Burns, Texas Tech University
Director of Education	Corine T. Norgaard, University of Connecticut
Director of Research	Nicholas Dopuch, Washington University
Director of Publications	Kenneth J. Euske, Naval Postgraduate School
Past President	Alvin A. Arens, Michigan State University

The Executive Director of the Association is Paul L. Gerhardt.

Membership in the Association is available at the following annual rates:

Full membership, US and Canada:	US $65.00
Full membership, other countries:	US $65.00
Associate membership:	US $12.00 for students residing in the United States and Canada

p 6243

In addition, full members may belong to one or more Sections: Auditing; Information Systems/Management Advisory Services; Two-Year Colleges; International Accounting; Public Interest; Government and Nonprofit; Taxation; Administrators of Accounting Programs; Accounting, Behavior and Organizations; and Management Accounting. Each Section assesses its own annual dues which are collected by the Association.

The Accounting Review is published quarterly by the American Accounting Association—January, April, July, and October: printed by George Banta Company, Inc., Menasha, Wisconsin. Subscription price is $90.00 per year.

Second class postage paid at Sarasota, Florida and at additional mailing offices.

Copyright © American Accounting Association. All rights reserved. Printed in the United States of America. ISSN 0001-4826.

Editorial correspondence and all manuscripts should be sent to Professor A. Rashad Abdel-khalik, Editor, *The Accounting Review*, Fisher School of Accounting, College of Business Administration, University of Florida, Gainesville, FL 32611.

The submission fee of $50.00 for members of the AAA or $100.00 for nonmembers must accompany manuscripts when submitted. Books for review should be sent to Professor Shane R. Moriarity, School of Accounting, University of Oklahoma, Norman, OK 73019. Request for reprint permission, subscriptions, advertisements, and dues should be sent to Mr. Paul Gerhardt, American Accounting Association, 5717 Bessie Drive, Sarasota, FL 34233. Notice of change of address should be mailed at least eight weeks in advance to ensure undelayed receipt of *The Accounting Review*.

THE ACCOUNTING REVIEW

A QUARTERLY JOURNAL OF THE
AMERICAN ACCOUNTING ASSOCIATION

Editor
A. Rashad Abdel-khalik

Associate Editors
Robert H. Ashton
Andrew D. Bailey, Jr.
Bala V. Balachandran
Victor L. Bernard
Ronald W. Hilton
John S. Hughes
Richard A. Lambert
James C. McKeown
Ira Solomon
Shyam Sunder

⑦

VOL. 67 JANUARY 1992 No. 1

HERBERT F. TAGGART, NORTON M. BEDFORD, R. J. CHAMBERS, SIDNEY DAVIDSON, CARL THOMAS DEVINE, THOMAS R. DYCKMAN, YUJI IJIRI, THOMAS F. KELLER, and ROBERT G. NICHOLS
A Tribute to William A. Paton 1

DAVID A. GUENTHER
Taxes and Organizational Form: A Comparison of Corporations and Master Limited Partnerships 17

DAVID N. RICCHIUTE
Working-Paper Order Effects and Auditors' Going-Concern Decisions 46

LISA KOONCE
Explanation and Counterexplanation During Audit Analytical Review 59

STEVEN F. CAHAN
The Effect of Antitrust Investigations on Discretionary Accruals: A Refined Test of the Political-Cost Hypothesis 77

A FORUM ON LABORATORY MARKETS AND AUDITING RESEARCH

NICHOLAS DOPUCH and RONALD R. KING
Negligence Versus Strict Liability Regimes in Auditing: An Experimental Investigation 97

DAVID E. WALLIN
Legal Recourse and the Demand for Auditing 121

THE ARIZONA STATE UNIVERSITY RESEARCH WORKSHOP
JAMES R. BOATSMAN, LAWRENCE P. GRASSO, MICHAEL B. ORMISTON, and J. HAL RENEAU
A Perspective on the Use of Laboratory Market Experimentation in Auditing Research 148

THE UNIVERSITY OF IOWA ACCOUNTING WORKSHOP
DOUGLAS V. DEJONG and ROBERT FORSYTHE
A Perspective on the Use of Laboratory Market Experimentation in Auditing Research 157

Contents continued on Overleaf

SHORTER ARTICLES

JENNIFER J. GAVER, KENNETH M. GAVER, and GEORGE P. BATTISTEL
The Stock Market Reaction to Performance Plan Adoptions 172

ASHIQ ALI, APRIL KLEIN, and JAMES ROSENFELD
Analysts' Use of Information about Permanent and Transitory Earnings Components in Forecasting Annual EPS 183

MICHAEL W. MAHER, PETER TIESSEN, ROBERT COLSON, and AMY J. BROMAN
Competition and Audit Fees 199

COMMENTS

WILLIAM M. CREADY
A Comment on the Proper Use of Residuals in Patell's Squared Unexpected Return Statistic 212

EARL K. STICE
A Reply to a Comment on the Proper Use of Residuals in Patell's Squared Unexpected Return Statistic 216

AMERICAN ACCOUNTING ASSOCIATION
Committees 1991–1992 219

BOOK REVIEWS, Shane Moriarity, Editor

Atkinson et al., ***Current Trends in Cost of Quality: Linking the Cost of Quality and Continuous Improvement*** .... WILLIAM R. PASEWARK 230

Banerjee, ***Contemporary Issues in Accounting Research*** ..... PARVEEN P. GUPTA 231

Innes and Mitchell, ***Activity Based Costing: A Review with Case Studies*** .... ALISTAIR M. PRESTON 232

Joye and Blayney, ***Cost and Management Accounting Practices in Australian Manufacturing Companies: Survey Results*** .... MARK A. COVALESKI 233

Macve and Jackson, ***'Marking to Market': Accounting for Marketable Securities in the Financial Services Industry*** .... R. D. NAIR 233

Ronen, Saunders, and Sondhi, ***Off-Balance Sheet Activities*** .... CHARLES W. MULFORD 235

Stevens, ***The Big Six: The Selling Out of America's Top Accounting Firms*** .... WILLIAM L. FELIX, JR. 236

White and Ohanesian, ***1991 Tax Penalties: The Complete Guide to Penalties Under the Internal Revenue Code*** .... DEBRA C. JETER 237

List of Books Received 238

Editorial Policy and Style Information 240

Occasional Reviewers 243

PAST EDITORS

1926–1929	William A. Paton	1968–1970	Charles H. Griffin
1929–1943	Eric L. Kohler	1971–1972	Eldon S. Hendriksen
1944–1947	A. C. Littleton	1973–1975	Thomas F. Keller
1948–1949	Robert L. Dixon	1976–1978	Don T. DeCoster
1950–1959	Frank P. Smith	1978–1983	Stephen A. Zeff
1960–1962	Robert K. Mautz	1983–1987	Gary L. Sundem
1963–1964	Lawrence L. Vance	1987–1989	William R. Kinney, Jr.
1965–1967	Wendell Trumbull		

THE ACCOUNTING REVIEW

A QUARTERLY JOURNAL OF THE
AMERICAN ACCOUNTING ASSOCIATION

Editor
A. Rashad Abdel-khalik

Associate Editors

Robert H. Ashton
Andrew D. Bailey, Jr.
Bala V. Balachandran
Victor L. Bernard
Ronald W. Hilton
John S. Hughes
James C. McKeown
Ira Solomon
Shyam Sunder

VOL. 67 APRIL 1992 No. 2

RAJENDRA P. SRIVASTAVA and GLENN R. SHAFER
Belief-Function Formulas for Audit Risk 249

STEVEN M. FLORY, THOMAS J. PHILLIPS, JR., R. ERIC REIDENBACH, and DONALD P. ROBIN
A Multidimensional Analysis of Selected Ethical Issues in Accounting 284

D. C. KIM
Risk Preferences in Participative Budgeting 303

A FORUM ON LIFO CHOICE OF INVENTORY VALUATION

SASSON BAR-YOSEF and PRADYOT K. SEN
On Optimal Choice of Inventory Accounting Method 320

ROSS JENNINGS, DAVID P. MEST, and ROBERT B. THOMPSON, II
Investor Reaction to Disclosures of 1974–75 LIFO Adoption Decisions .. 337

BARRY E. CUSHING and MARC J. LeCLERE
Evidence on the Determinants of Inventory Accounting Policy Choice 355

SHORTER ARTICLES

CHARLES D. BAILEY and EDWARD V. McINTYRE
Some Evidence on the Nature of Relearning Curves 368

STEPHEN K. ASARE
The Auditor's Going-Concern Decision: Interaction of Task Variables and the Sequential Processing of Evidence 379

DAN GIVOLY and CARLA HAYN
The Valuation of the Deferred Tax Liability: Evidence from the Stock Market 394

RICHARD P. BRIEF and RAEF A. LAWSON
The Role of the Accounting Rate of Return in Financial Statement Analysis 411

EUGENE A. IMHOFF, JR., and GERALD J. LOBO
The Effect of *Ex Ante* Earnings Uncertainty on Earnings Response Coefficients 427

Contents continued on Overleaf

COMMENTS

JOSEPH K. CHEUNG and MANDY LI
A Comment on "The Valuation of R&D Firms with R&D Limited Partnerships" **440**

TERRY SHEVLIN
A Reply to "A Comment on 'The Valuation of R&D Firms with R&D Limited Partnerships' " **443**

BOOK REVIEWS, Shane Moriarity, Editor

Curtis, ***Management Rediscovered: How Companies Can Escape the Numbers Trap*** .... FRED A. THORNTON, JR. **446**
Flesher, ***The Institute of Internal Auditors: 50 Years of Progress Through Sharing*** .... ARNOLD SCHNEIDER **447**
Fletcher, ***Government Accounting*** .... DAVID R. L. GABHART **448**
Graves, ***The Costing Heritage: Studies in Honor of S. Paul Garner*** .... MILTON F. USRY **449**
Gray, Roberts, and Gordon, ***Making Corporate Reports Valuable: A Survey of Corporate Reporting Practices by Major U.K. Companies*** .... ALAN J. RICHARDSON **449**
Harr and Godfrey, ***Private Sector Performance Measures and Their Applicability to Government Operations*** .... MARC A. RUBIN **450**
Kanatsu, ***TQC for Accounting: A New Role in Companywide Improvement*** .... WAYNE J. MORSE **451**
Martin and Kensinger, ***Exploring the Controversy Over Corporate Restructuring*** .... JANET I. KIMBRELL **452**
Maskel, ***Performance Measurement for World Class Manufacturing: A Model for American Companies*** .... RONALD V. HARTLEY **453**
McDowell, ***Ethical Conduct and the Professional's Dilemma: Choosing Between Service and Success*** .... HEIDI HYLTON MEIER **454**
Nunnally, Plath, and Johns, ***Corporate Lease Analysis: A Guide to Concepts and Evaluation*** .... DAN PALMON **455**
List of Books Received **456**
Editorial Policy and Style Information **457**

PAST EDITORS

1926–1929	William A. Paton	1968–1970	Charles H. Griffin
1929–1943	Eric L. Kohler	1971–1972	Eldon S. Hendriksen
1944–1947	A. C. Littleton	1973–1975	Thomas F. Keller
1948–1949	Robert L. Dixon	1976–1978	Don T. DeCoster
1950–1959	Frank P. Smith	1978–1983	Stephen A. Zeff
1960–1962	Robert K. Mautz	1983–1987	Gary L. Sundem
1963–1964	Lawrence L. Vance	1987–1989	William R. Kinney, Jr.
1965–1967	Wendell Trumbull		

William A. Paton
(1889–1991)

This issue of *The Accounting Review*
is dedicated to
the memory of
William A. Paton
(1889–1991)

THE ACCOUNTING REVIEW
Vol. 67, No. 1
January 1992
pp. 1–16

A TRIBUTE TO William A. Paton

PROFESSOR William Andrew Paton was the founder of *The Accounting Review* and was its first Editor (1926–1929). After a long illustrious career, Professor Paton passed away on April 26, 1991. Over the years, he has been the subject of several studies and undoubtedly will be the subject of many more. To show appreciation for Professor Paton's legacy, I have dedicated this issue of *The Accounting Review* to his memory and invited some of his former students and colleagues to share in paying tribute to this giant scholar. This tribute is a collection of short essays, each signed by its author. Except for the short biography by the late Herbert F. Taggart, all are invited and are written specially for this purpose. Each essay reflects some dimensions of Professor Paton that impressed the writer. I want to thank all the authors for contributing to this tribute and to extend my apologies to others who wanted, but were unable, to share in this occasion because time and space were limiting.

THE EDITOR
September 1991

(I am grateful to Harold Arnett for providing me with a copy of the short biography written by the late Herbert Taggart in 1964.)

WILLIAM Andrew Paton was born at Calumet, Michigan, on July 19, 1889. His father, Andrew Paton, who was superintendent of the nearby Osceola schools, had come from Scotland in 1851 at the age of nine months. His mother, born Mary Nowlin, came from a family of Yankees who got their start in this country before the Revolution.

Paton's early boyhood was followed by a year in Wisconsin, where his father was principal of Endeavor Academy, and ten years on a farm in the "mucklands" nine miles from Imlay City, Michigan. The latter move came in the spring of 1897.

Life was rugged in those years on the farm. "I recall," says Paton, "that for one year our total sales amounted to only $400, and from that had to be deducted taxes, insurance, wages for the hired man, etc. Not a big living for a family of six, even when the low prices of those days are taken into account. However, we were all hard workers and it never occurred to us, or to anyone else in the neighborhood, that we were candidates for relief by government or other outside agency."

Young Paton's formal schooling up to 1906 was pretty sketchy. He was a voracious reader, however, and the reading material on the family bookshelves was greatly augmented by the fact that his mother was the local custodian of the "state traveling library." This arrangement provided 50 assorted books from the Michigan state library each six months, and young William felt that it was his duty to see to it that no book went back to Lansing unread. He attended Imlay City High School for one year and spent another year at Michigan State Normal College (now Eastern Michigan University), but it was not until 1911 that he started seriously on a college program. By waiting on tables and washing dishes at boarding houses, by husking corn, sawing wood, and selling books from house to house, he managed to keep going and received an A.B. degree from the University of Michigan in 1915. The bachelor's degree would have come a semester or two earlier, but the $10 then charged as a diploma fee was too much for Paton's slender pocketbook.

Paton lost no time in completing requirements for the M.A. (1916) and the Ph.D. (1917). His doctoral dissertation was published by Ronald Press in 1922 under the title *Accounting Theory With Special Reference to the Corporate Enterprise*, and was republished 40 years later.

The teacher to whom Paton owed the most, and who was his chief inspiration, was Fred M. Taylor, a kindly but exacting master of logic and an economist of the neoclassical school. Accounting textbooks in those days were few and, for the most part, of poor quality. A recognition of the importance of accounting and the paucity of adequate material for accounting instruction probably furnished the challenge. In any event, Paton and his friend Russell A. Stevenson, a fellow instructor and graduate student, brought out the first edition of *Principles of Accounting* in 1916.

Paton's teaching career began in the the fall of 1915, when he was appointed an instructor in economics. He had, however, been a teaching fellow during his senior year. Earlier, in the summer of 1914, he had married Mary Sleato. The school year 1916–17 was spent at the University of Minnesota, but he returned to Michigan as assistant professor in 1917. The year 1918–19 found him in Washington, D.C., where he was first with the War Trade Board and later with the Income Tax Unit of the Bureau of Internal Revenue. An associate professorship brought him back to Ann Arbor in the fall of 1919, and he became a professor in 1921.

Throughout his active connection with the Michigan faculty, Paton retained his appointment as a member of the Economics Department, retiring in 1959 in possession of two titles: Professor of Economics and Edwin Francis Gay University Professor of Accounting, the latter being a special chair to which he was appointed in 1947. Economics had always been his second love, if not really, in a sense, his first.

The thousands of students who came under Paton's influence remember him as a stimulating teacher and a proliferate writer of textbooks. No student who encountered him has forgotton his incisive mind, his impatience with sloppy reasoning, his endless stock of homely anecdotes and Biblical references, his interest in his students' careers, and his amazing ability to remember names and faces.

He was one of the founders of the American Association of University Instructors in Accounting and became its sixth president in 1922. He midwifed *The Accounting Review*, the first issue of which appeared in March 1926, and nursed it through its first three years. In 1935, he was a ringleader in the move to drop the old, unwieldy name in favor of American Accounting Association, and to expand the association's functions and broaden its base of membership. Early in 1936, he was appointed the first research director of the newly designated organization, and in June, 1936, the *Review* carried the first product of the new research activity. This was modestly entitled, "A Tentative Statement of Accounting Principles Affecting Corporate Reports," and there are those who think that its appearance stimulated the American Institute of Accountants to organize the Committee on Accounting Procedure in an endeavor to make sure that upstart academicians did not monopolize the business of formulating accounting principles.

When the Committee on Accounting Procedure was initiated in 1939, Paton was a charter member, and he served for 11 years. His tenure on the committee was notable for his many dissents to the committee's pronouncements. He was not impressed with the sanctity conferred on an accounting method merely by long usage, he refused to kowtow to immediate expediency (though fond of insisting that expediency and logic are seldom at odds), he was a persistent foe of pussyfooting, and he was a poor compromiser. During Paton's membership, the committee adopted 40 of the 51 "Accounting Research Bulletins" which it promulgated in its 20-year existence. Though it is not possible to pinpoint the results of Paton's influence, they must have been substantial.

Paton's first practical experience in accounting came during vacations in 1914 and 1915, when he assisted in an investigation of the early income history of several Michigan electric power companies, and he maintained a strong interest in the field of public utility valuation and rate determination ever after. In 1927, largely as a result of the urging of his friend D. W. Springer (long-time secretary of the Michigan State Board of Accountancy), he took the C.P.A. examination and obtained a Michigan certificate. With F. E. Ross, another member of the Michigan faculty, he founded the public accounting firm of Paton and Ross, but his participation in the ordinary engagements of the firm was minimal, and he soon dropped out. He continued, however, to undertake consulting engagements from time to time, as his teaching schedule permitted.

Many honors came to this man. He was, of course, a member of all of the proper Greek-letter honoraries: Phi Beta Kappa, Phi Kappa Phi, Beta Gamma Sigma, and Beta Alpha Psi. He was a Fellow of the American Academy of Arts and Sciences. He was Dickinson Lecturer in Accounting at Harvard University in 1940. In 1944, Lehigh University gave him the honorary degree of Litt.D. When Ohio State University inaugurated its Accounting Hall of Fame in 1950, Paton was one of the first three initiates, and the only academician. His special professorial chair at the University of Michigan has already been mentioned. It was one of 12 named chairs provided by the Board of Regents to honor outstanding members of the faculty. Though this was not precisely an honor, he had reason to be proud of five years of service on the Ann Arbor city council, were he was a foe of waste and a champion of municipal efficiency.

Following "retirement," Paton taught at the University of Chicago, Olivet College (of which he was a trustee), Alma College, the University of Florida, and Wayne State University. He continued with occasional consulting work, with extensive writing, and some speaking engagements. He then had a little more time for his grandchildren and his hobbies. One of the latter is reflected in an article called "On Going Underground," in which he points out the advantages—not all connected with atomic warfare—of occupying dwellings beneath the surface of the earth. If this aberrant idea be taken by some of his critics as proof that his notions of economics and politics are those of a caveman, he would be totally unconcerned.

HERBERT F. TAGGART

Abridged and adapted by the Editor (with permission) from *Paton on Accounting*, Bureau of Business Research, University of Michigan, 1964.

I FIRST saw Bill Paton in 1947. He was addressing a group of management accountants and constantly moving his red head while clearly, critically, and analytically explaining without any doubt the need for accounting to develop more relevant information. The spellbound audience thundered its applause in appreciation, and I joined a multitude in adopting his thoughts as a beacon light for developing my own professional life.

Never smoking nor drinking, and quick to tell anyone how he felt about both, Paton applied his energies and considerable abilities to foster the establishment of an economics foundation for accounting theory and practice. Widely respected by accounting practitioners and academics alike, he used his influence to contribute significantly to the development of public accounting standards. Appreciative of his effective 11 years as a member of the AICPA Committee on Accounting Procedures, many in accounting practice credit it with the establishment of an objective, factual reasoning process for public accounting standard setting that carries over to this day.

A strong supporter of economic concepts as a fruitful base for the development of accounting standards, Paton frequently communicated his belief in the need to disclose economic reality by "poking fun" at accounting practices that reflected no economic substance. His horrified, "Was it fully depreciated?" exemplified his concern that accounting might recognize no loss when a valuable economic resource was destroyed. But political reality was more important to him than the routine use of economic concepts.

When public accounting faced massive criticism following the 1929 stock market crash, Paton subsequently joined with A. C. Littleton in classic support for the use of the matching concept as a means for measuring periodic income as a long-term success indicator. That shift in the emphasis of the accounting disclosures from the balance sheet to the income statement represented a fundamental change in accounting. No longer were users to expect that the accounting balance sheet accurately measured the current values of balance sheet assets. Rather, assets were to be considered as future expenses (Deferred Costs) to be matched against future revenues. The Paton and Littleton monograph went even further and encouraged users to direct their attention to the long-term success of an enterprise by deemphasizing the accounting concern for short-term fluctuating values in the balance sheet and emphasizing the information content of an income statement based on the matching concept.

The professional influence of the two authors was so substantial that their viewpoint overrode alternative proposals for adapting accounting to the post-depression era, and the accounting practice quickly accepted their *Introduction to Corporate Accounting Standards*. Accounting information users adjusted their views and began to treat the accounting periodic income measure as a relative measure of an organization's success rather than as an absolute measure of the increase in the current market value of the organization's resources and obligations. Conceptual and technological developments in the past 50 years, in the form of supplementary disclosures, risk measures, and portfolio balancing, may now provide means for disclosing current market values and external conditions in "Notes to financial reports." In the 1940s, however, the Paton and Littleton monograph did much to restore public confidence in the public accounting profession. So any tribute to Bill Paton is remiss to the extent that it neglects to recognize his contribution to all of economic society.

Influential as the Michigan professor was in improving the quality of accounting theory and practice, his contribution to education and to his students was equally significant. He passed on to his graduate students a pride and a goal that motivated them to contribute much to the development of improved information for economic society. On various occasions he had breakfast or lunch with former students at annual meetings of the American Accounting Association. He invited me to attend two such meetings, and I was astounded by his continuous efforts to inspire and encourage the further development of accounting thought.

Paton's textbooks emphasized why accounting practices were as they were or why they should be different than they were. His was no attempt to train students how to routinely apply current accounting rules and practices. He clearly sought to teach them how to think and develop relevant information for economic society. Successful as these textbooks were, and some suggest they could serve as models for developing today's accounting textbooks, the typical student's search for certainty in an uncertain world limited, to a degree, Paton's influence on accounting textbooks. To this day, too many textbooks train students in how to comply with current accounting rules, with little concern for alternative ways of developing more relevant and reliable information. Bill Paton, we still need you.

NORTON M. BEDFORD
University of Illinois at Urbana-Champaign

THE spirit of William Andrew Paton has long brooded over twentieth-century accounting. A prolific writer and speaker teacher of generations of students, many of whom followed in his academic footsteps, co-author of a monograph that must have had more influence over a longer period than any similar work—these are the accomplishments that are most commonly remembered. But Paton was a man of many parts. Through a period demanding that he play many roles, he played them all with characteristic bravura. I knew him personally only in his later years, through occasional encounters between 1959 and 1970. It was enough to evoke in me the same sentiments as those of his more knowledgeable eulogists, for I had immediate experience of his personal charm, his nimble wit, his abundant helpfulness, and his genial grace.

Born of teachers, later farmers through the economically chilly 1890s, Paton is said to have been familiar with the cut and thrust of hometown debate on economic and social matters. That setting has turned many adolescents into young radicals. But not Paton. Conservative by upbringing and disposition, he remained a lifelong champion of the opportunities of self-fulfillment through the free enterprise system, turning his pen in later life often in that vein to comment on current economic and social affairs. His early university years as a student, and most of his life as a teacher, bestrode the fields of economics and accounting. He was consultant to utility and industrial companies, appeared often before governmental committees, took a leading part in forming the American Accounting Association, and was the first editor of *The Accounting Review*. A master of disciplined writing and exponent of pungent prose, he gave piquancy to what is a generally unspiced and insipid literature. He loved people, loved audiences as actors are said to do, and they in turn drew from him his best in scintillating rhetoric. He thought nobly of his subject and its practice, serving the professional and academic communities with his abundant common sense, insight, and industry.

Countless thousands more have known Paton through his published work than through personal acquaintance. The crosscurrents of academic idealism and practical exigency, of private enterprise and public interest, of painstaking analysis and entertaining exposition, tugged his sympathies and their expression in different directions all from time to time. The effects deserve more attention than the work of such a colossus has attracted. Only two of his many publications can be noticed here.

In *Accounting Theory* (1922), Paton affirmed that "valuation is the crux of accounting," and, since values change, periodic statements should represent "current data [values] in terms of the actual dollar as of the date of the statements." If closer attention had been paid to developing the consequences of these ideas in the succeeding two generations, the profession may have been saved decades of fruitless debate, and the business community may have been saved decades of uninformed reliance on the products of mere conventions.

Paton's examination of the "postulates of accounting" in that work is a landmark exercise in the literature. The fundamentals of any professional art are crucial to its effective practice. Paton laid bare the unreality of the assumption of a stable money unit, the idea that costs attach, and other common assumptions. The struggle to shake oneself free of longstanding habits of thought has always been marked by audacious attempts to eliminate what seemed to be contradictory or unrealistic assumptions. *Accounting Theory* was a valiant adventure in that direction, in a field in which scholarly work was still in its infancy.

About 20 years later, the Paton and Littleton monograph appeared. *An Introduction to Corporate Accounting Standards* marked the apotheosis of the cost doctrine and the matching principle. That doctrine had its origin in the mid-nineteenth century, although Littleton's *Accounting Evolution to 1900* carried no hint of it, and Paton had described the balance sheet in 1922 as "the basis of the whole structure of financial accounts." Yet here in 1940 was a monograph without a single chapter on balance sheets, and the term doesn't even occur in the index. It has always been puzzling that, whereas the going-concern idea was noticed by Paton in 1922 and in the 1940 monograph, so little attention was paid in 1940 to the very measure that would minister to survival—balance sheets that would aid in the preservation of a viable relation between debt and the real debt-paying potential of assets. Paton's association with the venture is perhaps illustrative of his prodigal magnanimity, his readiness to serve the cause of accounting and the American Accounting Association, come what may. He may long have doubted the wisdom of his generosity. In any event, in 1970 he was moved to write: "For a long time I've wished that the Paton and Littleton monograph had never been written, or had gone out of print twenty-five years ago."

Perennially unassuming and diligent about matters of moment, he loved on occasions to jest. I recall vividly an after-dinner address at the annual meeting of the AAA about 30 years ago when he regaled a delighted audience with recollections of early days of the Association to gales of laughter and waves of applause, while his wife, Mary, tugged his coattails at his bland irreverence.

Always down-to-earth, he was not tolerant of pretentiousness, especially in his own bailiwick. "I must confess," he said on one occasion, "that I get very tired at times of what the current crop of revolutionaries in accounting have to say, especially those who find excuses for abandoning or avoiding the term 'accounting' entirely." That, from one who in his own time and way was a revolutionary. Paton was an accountant for all seasons.

Above all things, through a tumultuous century of changes in education and in commerce, Paton unceasingly thought greatly of his twin professions. That was the source and the mark of his greatness.

R. J. CHAMBERS
The University of Sydney

THE opportunity to write this tribute to Bill Paton brings back many stirring memories. Last year marked 50 years since I took my first accounting course from Bill at the University of Michigan. (At the party celebrating his hundredth birthday, he reminded me what seat I had occupied in that class. I responded that his memory was better than mine because he was so much younger than I, and in many ways he was even then.) In that course and the ones that followed, his enthusiasm, his probing mind and keen wit, his dedication and devotion to clear thinking in accounting were powerful enough to cause me to change all of my career plans. He won me over, as he did so many others, to a career in accounting. For half a century he has remained the single most important professional influence on my life.

Of course, that's not to say that we always agreed on political, social, and economic questions. From Franklin Roosevelt to Ronald Reagan, we had strong differences of opinion on the quality of presidential performance. Bill had strong opinions and never hesitated to voice them forcefully. However, on accounting questions, he has always been supreme. I, along with so many others, can never think of accounting without thinking of Bill Paton.

Throughout his career, Bill emphasized the role of accounting in economic analysis. He was delighted to have the *Trueblood Report* in 1973 adopt his views and say, "The basic objective of financial statements is to provide information useful for making economic decisions." When the FASB in 1978 repeated this view in almost identical words, it helped to reduce his concerns about the FASB's insistence, as he put it, in "producing cookbooks, not principles statements." In one of his last major public appearances at the 1983 meeting of the AAA in New Orleans, he gave his memorable dream-sequence speech about the FASB, brandishing his cane and demonstrating how much vigor and imagination a man of 94 years can have. Those of us who were present will never forget his performance. That speech demonstrated his keen analytic ability but also his skill at explaining a complex issue in a way that was clearly understood and colorful enough to remain in the listener's or reader's memory for years to come.

By a remarkable coincidence, the AICPA (or its predecessor body) was formed two years before Bill Paton was born. For much of the century of their joint existence, Bill Paton was a powerful influence on the AICPA's activity. In his 11 years on the Committee on Accounting Procedure and with his considerable stock of homely anecdotes and Biblical references expressed in many articles in the *Journal of Accountancy*, he helped to shape the profession's statements on GAAP. Bill regularly and forcefully urged the standard-setting bodies to recognize the importance of sound economic principles in official pronouncements and to be modest in offering edicts. In one exasperated article in the *Journal of Accountancy*, he said that "the Board [the APB] has no business laying down the law." As an indication of the esteem in which he was held in the profession, when the AICPA decided to honor the "Academic Accountant of the Century" at its centennial celebration, no name other than Bill Paton's was mentioned at the committee meeting where the selection was made.

For the thousands of students, many of whom became leaders in industry and in the profession, and for the hundreds of college teachers whose careers were shaped by him, there will never be anyone like Bill Paton. His memory will be with us always.

SIDNEY DAVIDSON
The University of Chicago

HERE is that rare Shavian anomaly—"a genius without ostentation"—no beards, no tweeds, no twirling keys, no academic verbiage, no computer jargon, no inside references to Kolmogorov, no recourse to volterra equations of the second kind, no hopping from one paradigmatic bandwagon to another.

The viewpoint remained stable, the framework of neoclassical economics remained firm, the value system—the virtues of hard work, personal and intellectual honesty, rational discourse, and the free market—never varied. Behavorial studies were interesting but more or less confirmed traditional intuitions. Market studies were interesting extensions of established views about the incorporation of new information. (How could anyone who ever paid a telephone bill believe that information is costless?) The newer contract-agency orientation with its ugly corporate bureaucrats is more depressing than his own view of corporate management. My own tentative attempt to treat accounting as literary discourse was interesting but aroused no great enthusiasm.

Yet this bandwagon resistance did not interfere with an occasional reappraisal of the financial environment and the need for revised professional guidelines. His commitment to current costs, for example, remained constant, but their importance diminished during the Depression, when prices were relatively stable, so that he could go along (reluctantly) with Littleton and the traditional locked-in, hold-to-maturity concept of historical costs. Strange bedfellows!

Plaintive comment: "Some critics call me stubborn and dogmatic while others call me wishy-washy." Both may have been right.

Stubborn? Stubborn enough for two Missouri mules. He tried to get it *right!* Wishy-washy? Sometimes. In the scientific tradition, he kept an open mind until the evidence seemed to support his belief. (Strangely enough, after his pioneering postulates, Paton lost interest in presenting accounting as a science even though at least two from his immediate family were internationally recognized physicists.) Argumentative? You bet! Now and then he may have misread the cues and on rare occasions he may have been downright wrong. Much of the time he argued straightforward preference statements (normative positions) with passion and overwhelming persuasion. Thus he was sometimes caught in that netherland where propositions are neither true nor false—divorced from sense content. Other statements were sometimes non-operational, e.g., with limited revenues all costs are recovered *pro rata*. Yet even his fiercest antagonists recognized a great mind and an intellectually honest man.

Needless to say, his sharp mental reactions, powerful rhetoric, and dominating personality were intimidating to ordinary doctoral candidates. On occasion he thundered like an unbearded prophet—a strident voice from the Hiawathan wilderness—at other times he listened to struggling students with complete attention. He did not suffer fools and laggards with noticeable grace, but he was a firm supporter of serious students, especially if attacked by overly aggressive faculty members or harassed by the petty requirements of early doctoral programs.

Assessments were honest and, except for an occasional reference to trash or junk, were seldom harsh. Comments (paraphrased here) were no nonsense, straight-from-the-shoulder with unmistakable meaning.

Regarding a dissertation submission: We can't expect major doctoral contributions and we sometimes apply a modified labor theory of value. This effort is pretty rough stuff, but you did pitch in and get *something* done.

Regarding a not-too-subtle inquiry about his reluctance to cite historical precedents: Most of what I say is common knowledge. Certainly I've never meant to imply that all of our ancestors were damn fools.

Regarding an over-exuberant criticism of a leading public accountant: George is a plodder and may be a little light upstairs, but I'm not sure that you have earned the right to be his executioner.

Regarding CEOs: At retirement roll my chair into the sunshine and give me a lapful of annual reports. The dubious excuses and gracious bows will help the days go by.

A half century or so ago, Professor Paton remarked wistfully that he had spent a *lifetime* of hard work and at best had made only a few minor scratches on the surface of accounting knowledge. Clearly his colleagues and students take a different view and insist that he destroyed at least some of the arcane tenets of the ancient accounting guild. How deep are these scratches and how may they be characterized?

A contribution of some importance resulted from his relentless campaign to clean up the semantics of the profession and to improve its terminology. Individual recommendations were often minor, e.g., removal of such rummage as surplus, reserves, profit and loss statement, gross profit. Yet general adoption of these simple recommendations required a half century of constant persuasion.

Structural changes were chiefly in the direction of forming a consistent structure, and here again specific recommendations were not world-shaking. Even the entity orientation was already well established, and his contribution was primarily in his messianic insistence that it was the appropriate organizational structure for an industrialized world.

The teaching tradition, with its literate texts and diabolical problems, was an important contribution. Paton was a great educator, and over the years a small army of dedicated and energetic young disciples went forth to spread the Patonian unorthodoxy and make their own small contributions.

Paton's insistence on the rigorous application of economic theory to reasoning in accounting was a shrewd insight and a major contribution to the field. Accounting, to Paton, is an extension of economics to the operation of the firm, and the accounting entity is an organization—an institutional arrangement to produce economic goods. Yet the world agrees that his most important contribution came from his respect for logical thinking, his abhorrence of shoddy argument, and his unrelenting commitment to intellectual honesty.

With his leadership, accounting became, during his lifetime, a genuine academic discipline worthy of university study and a legitimate member of the social science establishment. What an impressive human being, William Andrew Paton! What an opportunity to have studied with him! What a joy to have known him!

CARL THOMAS DEVINE
Florida State University

IT was August 1979, and as I walked up to the Sheraton Waikiki Hotel in Honolulu, I was assailed by an impeccably dressed gentleman who had apparently lost his luggage. For several minutes he confided to me what was wrong with the younger generation and, in particular, those placed in charge of his (and others') suitcases. Finishing by reminding me, as he had done in years past, of the evils of drugs and smoking, he turned and moved quickly away into the hotel, apparently realizing I would be of little help in retrieving his missing possessions.

I was to be in his company one more time while in the islands. The occasion was the annual luncheon for the past presidents of the American Accounting Association. As a member of the Executive Committee, I was invited to attend along with those past presidents who had made the long trip to Honolulu.

When we took our seats for lunch, the head of the main table was reserved for the Association's Elder Statesman and Past President in the year 1922, William Andrew Paton. He was in excellent spirits, apparently having recovered his missing suitcases, and proceeded to tour the room greeting everyone, including myself, by name (despite attempts by colleagues to intercede with introductions) and asking a personal question on some facet of our lives that could only have come from his exceptional memory.

I heard little of the luncheon conversation because I had been relegated, as was due my station, to a smallish table in a far corner. I was told later, however, that at one point Professor Paton was reminded of his famous monograph written with Professor Littleton. "Ah yes," Paton was reported to have said, "Paton and Littleton, profit and loss. I was the profit. . . ." To my knowledge the sentence was not completed. As lunch concluded, however, a partner of one of the major firms and a student of Bill's in the mid-thirties was ushered into the room. Bill, obviously pleased and surprised to see him, immediately greeted him by name and asked after his family. Now, to say the least, this rather unnerved our visitor, who nevertheless launched into a brief tribute to his mentor. It is what came next that I have never forgotten and want to share with you.

"Bill," the speaker said to the assembled guests, "was a demanding teacher. In fact, he gave me a grade of C in his accounting course." At this point, Bill turned abruptly in his chair to look up at the speaker, waggled a finger under his nose, and declared, "It was not a C, it was a C minus!" "No Bill," our speaker retorted, "it was a C. I remember because it was the only C I ever received." "You're wrong, young man, it was a C minus. You failed to solve the consolidation question on the final correctly, and I was forced to lower your grade." "Well, I guess you must be right," demurred the perplexed ex-student.

So what is so important about this simple event in the life of a man who at that time was 90 years of age, a man who had accomplished about all there was to accomplish in his chosen field? Indeed, it is precisely because of his stature that the event is significant. The process of education, teaching, and, most of all, his students were the center of Bill Paton's professional career. If we learn one thing from this man, it is that our students are the most precious gift we are allowed. Each is to be cherished and if possible given a special place in our lives. Surely this was true for Bill.

THOMAS R. DYCKMAN
Cornell University

IN the fall of 1979, Thomas J. Burns, then Book Review Editor of *The Accounting Review*, asked me to review the classic monograph, *An Introduction to Corporate Accounting Standards* by William A. Paton and A. C. Littleton, in commemoration of its fortieth anniversary. For some reason, the monograph had never been officially reviewed in *The Accounting Review*, in spite of the significant influence it had over many scholars, practitioners, and students. I undertook the task of preparing the review with trepidation. After all, a book that has been already so widely read is much more difficult to review than a book that was just published.

After months of work, I sent a copy of my review article to Professor Paton. Upon careful review, he requested to have his statement published along with the review article, which appeared in the October 1980 issue of *The Accounting Review*.

In the statement, after providing additional insight into the historical perspectives surrounding the monograph, he concluded that "Perhaps the main reason I am writing this statement is to make my personal stand clear. I have always been a *value* man. This is partly the result, no doubt, of my training in neo-classical economics, and several years of teaching in that area before wandering into accounting."

Yet, the monograph is, with the exception of a final section, almost exclusively cost-based. Only after the first six sections, where a logical construction of essential ingredients of accounting is presented, does it discuss replacement cost disclosure in the footnotes to financial statements.

It had long been a puzzle to me why Bill Paton, whose earlier writings advocated the recognition of appreciation in income so strongly, agreed to make the first six chapters of the monograph so exclusively cost-oriented. Of course, influence from his co-author, A. C. Littleton, a strong advocate of the historical cost principle as evidenced by his writings, as well as the members of the Executive Committee who worked closely with the two authors, must be taken into account. I felt, however, Bill Paton must have seen some merit in presenting the materials in this way.

The puzzle was partially resolved when I read an excellent paper by Stephen A. Zeff, "Paton on the Effects of Changing Prices on Accounting: 1916–55," in *Essays in Honor of William A. Paton* edited by him and his colleagues in 1979, in which he carefully depicted the transition of Bill Paton's position on the effects of changing prices over four periods.

It seems that the monograph was written during the third period, when Bill Paton's recognition of the importance of the "accountability side" of accounting was strongest, even if his heart was still on the "decision side" of accounting.

Accounting seems to contain two sides that are almost exactly opposite of each other in characteristics. They may be called the "decision side" of accounting and the "accountability side" of accounting. The former is centered on the objective of providing information useful for decision making, while the latter is centered on the objective of providing information on charges and discharges of accountability so that accountability relations can function smoothly. The two sides can be characterized by such contrasting terms as *freedom vs. responsibility*.

Decision making means a choice among alternatives, and a choice means freedom. Decision making is an "intrapersonal" activity that can be made alone as a free person. Accountability, on the other hand, is an "interpersonal" activity that cannot be made alone because it involves two parties, an accountor and an accountee. An accountor accounts for his or her activity for the benefit of an accountee.

The two sides of accounting are built on fundamentally different bases, yet they are connected as two sides of a coin. This is because freedom and responsibility are linked together in most (or, as some people think, in every one) of our decisions. Especially in modern management where a manager makes a decision on the use of someone else's money, every decision has accountability attached to it.

It seems very fortunate for the readers of the monograph that the two authors, who came from different perspectives on accounting, were able to agree on the underlying orientation of the monograph. The structure they adopted is an excellent one.

After completing the monograph with A. C. Littleton, Bill Paton became a feverish campaigner for replacement cost and price-level adjustment. Having been exposed to the accountability side as well as the decision side of accounting, Bill Paton must have become a persuasive campaigner far more profound than he was in 1916–18, the first period "typified by youthful idealism" according to Steve Zeff.

In spite of the apparently drastic swing in his position, Bill Paton seems to have been quite satisfied, according to his statement, with the contents of the monograph aside from a few minor changes and additions he might have made. However, he wished to make it clear that he had been a *value* man all along. Here, we may see a mild detachment between a book and its author. But this is the way it should be since, after all, a book is the author's intellectual child, with a life separate from the author.

Professor Paton's longevity has benefited the members of the AAA greatly. And we all know that he will be remembered through his many superb intellectual children, whose longevity will extend many centuries into the future. *An Introduction to Corporate Accounting Standards* is only one of them.

YUJI IJIRI
Carnegie Mellon University

(The author acknowledges the helpful comments of Stephen A. Zeff.)

MY first acquaintance with Bill Paton came via letter in response to my application for admission to the MBA program. He began the dialogue with the assertion that Michigan was the best place to study accounting and furthermore that I should pursue the Ph.D. and become a teacher and researcher.

Upon arrival at the University, I went to his office to meet him and to discuss my course of study. His opinion was clear and concise—I should study asset accounting and corporate financial statements; after this, I should take his seminar on the theory of income measurement, do what was required for the degree, and complete the elective set with those subjects that might be of interest.

I had studied accounting as an undergraduate and had practiced as an auditor, but I had never understood accounting theory, Patonian-style, before. Bill Paton at age 66 had strong opinions on many subjects, especially the role of government in our society, but he had a way of explaining accounting that was very effective. His discussions of why entities buy goods and services and how they use these resources to produce products or services to generate revenue made accounting understandable as a measurement process of an ongoing activity. The rules, of which there were very few (mainly ARB No. 43 and a small monograph by Paton and Littleton), were held in low esteem by Bill Paton. I often wonder what he would say about our attempts to teach the volumes of rules promulgated by the APB and FASB.

As we all know, Bill Paton had little regard for academic material or issues that he viewed as softheaded. He clearly viewed the proper role of government in terms of the least government as the best government. The individual was all-important to Bill Paton. His "shirt sleeve economics" espoused a philosophy for which Paton was famous—at his retirement dinner in 1959, he was offering his customary advice that

Social Security was a curse on our society and that television was a waste of time that should be used to exercise the brain and continue the learning process.

Needless to say, Bill believed in the market system. One day at lunch in the Michigan Union when he was confronted with apple pie as the only selection, he asked why there were no choices. The attendant replied, "Dr. Paton, what would you like? I will get it for you." Bill responded, "I want apple pie, but I want a choice."

Bill never thought much of academic administrators, and shortly after I became Dean at Duke, he told me that perhaps I should lend a helping hand to the institution, but he cautioned that the tenure should be short, lest I, like others, become a bureaucrat. When I last saw Bill, he again cautioned that in spite of any success as a dean, I should return to teaching—maybe, there is still time, Bill!

Bill Paton's infectious smile, his commitment to his students, and his love of his work have been an inspiration to many who left his coaching to grow into positions of major responsibility. Those with whom I have had some acquaintance always speak of Bill as a true friend and a person for whom they have great respect. Well after retirement, Bill Paton had a strong mind, a conviction about economics, and a deep interest in those with whom he had worked. We will miss his straight thinking and lack of tolerance for softheaded dialogue. Anyone who studied accounting with Bill Paton will have been forced to think through the mechanics of the double-entry procedures to understand the economic phenomena that are reflected in the numbers. Bill Paton's influence on the development of accounting theory and thought is without peer. Those who had the privilege of studying with Bill and becoming his friend will never be able to view accounting as the process of applying rules. They will instead insist on understanding the economic events and on the use of the measurement and disclosure concepts that provide the most accurate representation of the facts.

Let us hope that the Patonian approach is not dead, but can be perpetuated and may yet prevail. Bill Paton, with his logic and his philosophy, stands as a giant among us. I will never forget his smile, his red hair, or his approach to teaching accounting.

THOMAS F. KELLER
Duke University

A FEW years after graduating from the University of Michigan, I returned to Ann Arbor and decided to visit Professor Paton. Standing outside his classroom, I did not expect him to remember me, but as he exited the classroom he looked up and immediately called me by name and asked about my new career at Price Waterhouse. Professor Paton had a remarkable memory, but it was just one of the many distinguishing characteristics of a man I will never forget.

Over the next 30 years, I saw Professor Paton many times, and his photograph has occupied a prominent place in my office—a constant reminder of his immense influence on my career. Of course, it was not just my career that was affected. Even though he was close to retirement, Professor Paton taught three courses in the early 1950s—two of which were undergraduate courses. The thought of the leading accounting theoretician of this century teaching undergraduate courses and the related impression on these students is worth considering.

He was an exciting teacher with lovable characteristics and unlimited energy. In class, he gained our undivided attention with vivid language and endless anecdotes about his childhood on the farm. We had to pay close attention because we did not want to miss something really special. The sparkle in his eyes revealed how excited he was about the subject matter.

Professor Paton always challenged the established theories in accounting, even when he had helped establish them. Our textbooks were written by him, and he always challenged the status quo in them. In his lectures he would challenge his own textbooks. "Disregard Chaper 12," he would say, "my thinking on this subject has changed. Consider this lecture in lieu of it." Accounting was not a static subject to Professor Paton, and that added to the intellectual stimulation of the classroom and our interest in accounting as a profession. Professor Paton's classes emphasized the importance of the written word. He was a tough critic and a stickler for effective communication, regularly downgrading papers that were not well written. He was impatient with sloppy reasoning, whether on paper or in classroom discussion, and very sharp responses to poorly thought-out answers were more the rule than the exception.

Professor Paton was never reluctant to speak out. He tried to shape opinions. He held many strong opinions, and he was not bashful in sharing them. For example, I recall asking him if I should take a course in governmental accounting. "If you are really interested in such a dull subject, check out a book from the library and read it over the weekend. Don't waste an entire semester," he said. Years later, I recall him making fun of business school courses with names like "Organizational Behavior."

Professor Paton did not favor a vast array of rules governing every conceivable situation that might arise in the world of accounting. He believed common sense and sound judgment exercised by highly principled professionals was superior—a view shared by George O. May, a practitioner and contemporary with whom Professor Paton formed a close bond. He often referred to him in discussions with me. This view, like many of his other views, has stood the test of time.

When the American Accounting Association named Professor Paton the Accounting Educator of the Twentieth Century, I thought it was richly deserved. His legacies are the thousands of students who sat in his classroom and the other thousands who profited from his thoughtful textbooks. We are all the better for the experience.

ROBERT G. NICHOLS
Price Waterhouse

THE ACCOUNTING REVIEW
Vol. 67, No. 1
January 1992
pp. 17–45

Taxes and Organizational Form: A Comparison of Corporations and Master Limited Partnerships

David A. Guenther
University of Connecticut

SYNOPSIS AND INTRODUCTION: Much of the research in positive accounting theory deals with whether managers choose accounting methods that reduce or minimize certain costs faced by firms, such as the cost of violating bond covenant restrictions, political costs, and the tax cost associated with the use of the FIFO rather than the LIFO inventory method. However, this research ignores some more fundamental costs associated with the legal form under which the firm chooses to operate, a much larger issue that precedes the choice of accounting methods. This article examines this larger issue by focusing on the trade-off that exists between tax costs and transaction costs in the choice of organizational form.

Scholes and Wolfson (1986, 1989) assert that an organization's form is chosen to minimize both tax costs and transaction costs. Under this theory, if the corporate form has a greater tax cost than that of an alternative form, the partnership, the corporate form would not be chosen unless the transaction costs of the partnership form exceed those of the corporate form. Fama and Jensen (1983b, 327) examine costs of alternative organizational forms and state that "the form of organization that survives in an activity is the one that delivers the product demanded by customers at the lowest price while covering costs." Although Fama and Jensen specifically excluded tax costs from their analysis, Scholes and Wolfson argue that both relative tax costs and relative transaction costs are important determinants of organizational form choice, and that changes in relative tax costs will result in changes in form. They also make the following predictions

I would like to thank my dissertation committee, Lauren Kelly (chair), Jon Karpoff, and Terry Shevlin. Helpful comments on earlier versions of this article were received from workshop participants at the Universities of Arizona, Connecticut, Iowa, New Hampshire, Texas, and Washington, Duke University, Colorado State University, and Ohio State University. I would also like to thank the two anonymous referees for their comments. Financial support from the University of Connecticut and the Cox memorial scholarship fund at the University of Washington is gratefully acknowledged.

Submitted March 1991.
Accepted August 1991.

about the effect of the 1986 Tax Reform Act: corporations will be replaced as an organizational form by partnerships for new ventures financed with equity; many existing corporations will convert to partnership form; absent this conversion, many corporations will add debt to their capital structure.

These predictions lead directly to two questions. First, how do tax laws affect the choice of business entity? Second, what happens when these tax laws change? These two general questions are investigated here through two more specific research questions. The first is, What are the incremental tax costs and transaction costs of alternative organizational forms (corporations and master limited partnerships) available to large, publicly traded firms? The second is, How will managers of existing corporations respond to a tax law change (the 1981 Economic Recovery Tax Act) that causes the tax cost of the corporate form to increase relative to that of the partnership form?

The results of empirical tests of four hypotheses developed to investigate these research questions indicate that, for the period 1978–85, the corporate form resulted in a significantly greater average tax cost than the partnership form, and this incremental tax cost increased significantly after the 1981 Economic Recovery Tax Act (ERTA). Partnerships are found to have a significantly lower return on assets and sales than a matched sample of corporations. Responses of managers to increasing tax costs of the corporate form after 1981 are predicted to be (1) increasing long-term debt, (2) increasing non-dividend distributions, and (3) decreasing dividend payout ratios. Empirical results for both univariate and multivariate tests are consistent with these predictions.

These results make a significant contribution to accounting research for two reasons. First, they demonstrate a relationship between changes in relative tax costs of alternative (competing) organizational forms and changes in such fundamental elements of capital structure as debt levels, non-dividend distributions, and dividend payout ratios. Previous research on the relationship between taxes and capital structure (e.g., Bradley et al. 1984; DeAngelo and Masulis 1980; MacKie-Mason 1990; Titman and Wessels 1988) has not considered the effect of the tax cost of alternative organizational forms. Second, the results provide empirical support for predictions by Scholes and Wolfson (1986, 1989) and Petruzzi (1988) regarding responses of corporations to increases in relative tax costs (i.e., Scholes and Wolfson predict increasing long-term debt; Petruzzi predicts increasing non-dividend distributions).

Key Words: *Double taxation, Partnership, Organizational form, Dividend clienteles.*

Data Availability: *Data are available on request from the author.*

THE remainder of this article consists of four sections. Section I develops four hypotheses dealing with the relationship between tax costs, transaction costs, and organizational form. Section II discusses some measurement issues related to the variables used in the empirical tests, as well as the sample selection procedures. Section III reports the results of testing the hypotheses, and section IV concludes the discussion.

I. Development of Hypotheses

Hypothesis H1

Because of double taxation, the tax cost of the corporate form is predicted to be greater than that of the partnership form. Stated in null form, this hypothesis is:

H1: The incremental tax cost of the corporate form over the partnership form is less than or equal to zero.

As used here, the "tax cost" of an organizational form refers to the total of the income taxes imposed on the organization (e.g., the corporate income tax) and any taxes imposed on the equity owners of the organization (e.g., the shareholders' tax on dividend income). However, the presence of double taxation in the U.S. tax system is not a sufficient condition for the corporate form to always result in a greater tax cost than the partnership form. Whether the tax cost is greater for the corporate form (i.e., the incremental tax cost of the corporate form is positive) depends on the relative corporate and shareholder marginal tax rates, as well as the level of dividends. This relationship is illustrated in figure 1.

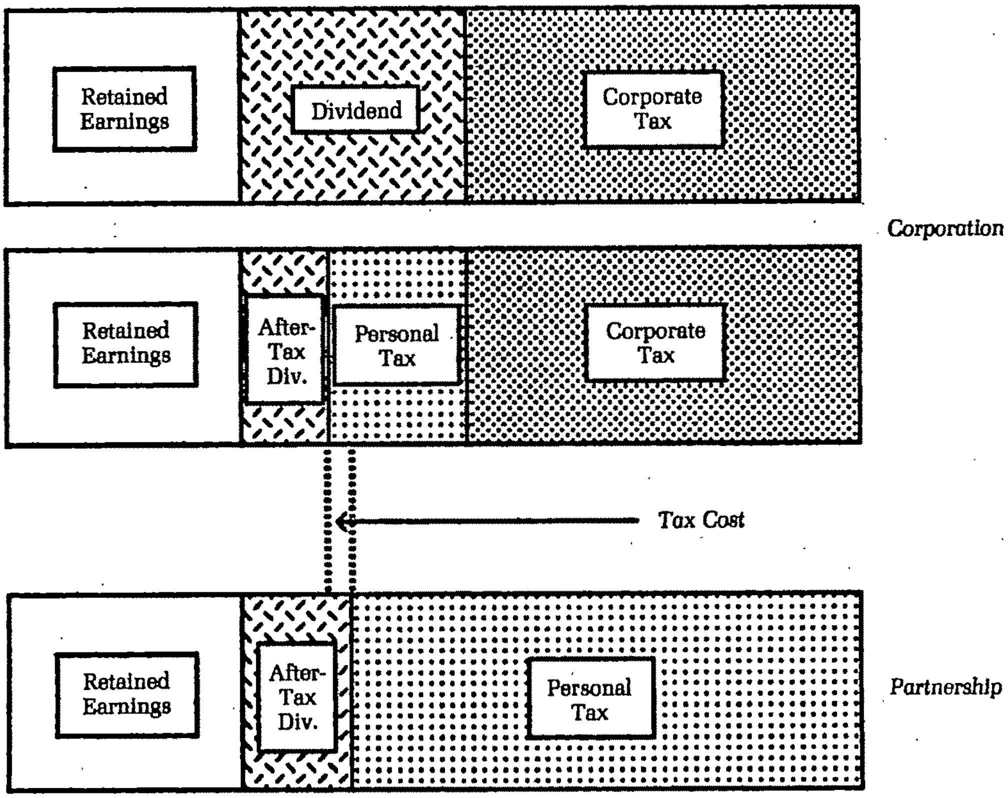
The top two diagrams in figure 1 illustrate a situation in which a corporation with a tax rate of 46 percent pays out 50 percent of its after-tax earnings as a dividend to shareholders whose tax rate is 60 percent. If this same corporation could costlessly change its form to a partnership, the bottom diagram illustrates the total personal income taxes to be paid by the partners, assuming that the pre-tax earnings are the same and that the partners have the same tax rate as the shareholders. Under the above assumptions, the corporate form incurs a tax cost relative to the partnership of $0.022 per $1 of income (the distance between the two dotted lines in figure 1), as follows:

Corporate tax: $1 × 0.46 =	$0.460
Shareholder tax: [0.50 × $(1 − 0.46)] × 0.60 =	0.162
Total tax, corporate form:	$0.622
Partnership tax: $1 × 0.60 =	$0.600
Difference (incremental tax cost)	$0.022

Hypothesis H2

Under the Scholes and Wolfson (1986, 1989) theory of organization choice, the form of organization is chosen to minimize tax costs and transaction costs. Prior to 1981, the partnership form of organization did not allow for publicly traded ownership interests, a form of transferability restriction that imposed a major transacton cost. However, after 1981 the partnership form offered both limited liability (through the use of a limited partnership) and publicly traded ownership interests (through the use of a master limited partnership). It is predicted (in hypotheses H3 below) that the incremental tax cost of the corporate form increased after 1981. Given this combination of a decrease in the transaction cost of the partnership and an increase in the incremental tax

Figure 1
Incremental Tax Cost of a Corporation Compared with a Partnership



Note: It is assumed that the corporation has a tax rate of 46 percent and distributes 50 percent of its after-tax income as a dividend to shareholders. The personal income tax rate for shareholders and partners is assumed to be 60 percent.

cost of the corporation, corporations would then survive as a potential form of business after 1981 only if there are additional transaction costs of the partnership form that exceed the incremental tax cost of the corporate form. If this were not the case, new businesses would be started as partnerships rather than corporations. Further (absent additional transaction costs such as taxes on liquidation), firms would maximize their value by changing from corporations to partnerships. This motivates the second hypothesis:

H2: The incremental transaction costs of the partnership form over the corporate form are less than or equal to zero.

If a firm can be characterized as a nexus of contracts (Jensen and Meckling 1976), then differences in transaction costs are expected to exist when there are differences in contractual relationships between partnerships and corporations. By examining those cases in which contracts entered into by master limited partnerships (MLPs) differ from those entered into by corporations, it is possible to identify potential transaction costs. For example, one significant difference between partnerships and corporations involves the lack of voting rights for limited partners. Although the general partner may be replaced through a vote of the limited partners, the effectiveness of this technique as a control mechanism is reduced by two constraints. First, to replace the general partner requires at least the vote of a majority of the limited partner units (and in many cases the percentage is substantially greater than a simple majority). Second, most partnership agreements give the general partner the ability to convert general partner units into limited partner units, which may then be voted by the general partner. In some cases this makes it literally impossible for the limited partners to ever replace the general partner. Another (anecdotal) example of greater transaction costs for MLPs is found in the 1988 prospectus of Apache Petroleum Company, an MLP whose management states that changing from a partnership to a corporation would save $4 million (approximately 2 percent of revenues) in annual overhead costs, such as preparing and providing complicated tax information to the limited partners.

Differences between partnerships and corporations are evident in contracts between the firm and the government (e.g., reporting and recordkeeping requirements), the government and the partners (e.g., tax-filing complexity), the firm and its general partner (e.g. unlimited general partner liability), the partners and managers (e.g., managers are often employees of the general partner rather than of the partnership), and the limited partners and the general partner (e.g., lack of participation in management through a board of directors, lack of voting rights). It is predicted that these differences all result in transaction costs to the partnership or its equity owners that exceed the transaction costs arising from the corporate form of organization.

The definition of transaction costs from Williamson (1981) and Arrow (1969) is the cost of running the economic system, or, for this study, the cost of running a corporation or partnership. In the context of organizational form choice, the decision maker is assumed to consider all costs of running the organization. If the total of all costs of the partnership form, regardless of the type of costs, exceed the total of all costs of the corporation form, the corporation will be chosen as the form of organization (assuming equal revenues). Hypothesis H1 focuses on tax costs, while hypothesis H2 focuses on all other costs, labelled transaction costs.

Transaction costs can be divided into two categories: explicit and implicit. Explicit costs are actual expenses incurred by the partnership, such as bookkeeping, recordkeeping, and data-processing costs. They also include the costs of perquisites consumed by opportunistic managers, monitoring and bonding costs to limit such consumption, and the premiums for insurance required by contracts to indemnify managers for any losses that result from their actions (as required in most partnership agreements).

Implicit costs are those resulting from managers making financing and investment decisions for the partnership that differ from those that would be made by the equity owners of the partnership (i.e., "agency costs." For a discussion of these costs in connection with oil and gas tax shelter limited partnerships, see Wolfson [1985]). Karpoff and Rice (1989) also include in this category the costs of incompetent or inefficient managers. Since implicit costs are not actual expenditures of the partnership, but rather are opportunity costs, they are not reflected directly in the cost of goods sold, or general and administrative expenses. However, these costs will affect the performance of the partnership, and partnerships with high implicit transaction costs will be less profitable and less efficient in utilizing their assets than those without such costs.

P6243

Note that costs of actually forming or dissolving an organization are not directly examined. Costs of forming the organization are expected to be reflected in higher general and administrative expense (e.g., "legal expense") in the year of organization, or in additional amortization expense of capitalized intangible assets in future years. Costs of dissolution are not addressed, since the cost of dissolving an existing corporation is assumed to be so great (because the liquidation is a taxable event) as to prevent a direct change in form for most firms.

Hypothesis H3

Hypothesis H3 examines the effect of the 1981 individual tax rate reductions on the incremental tax cost of the corporate form. From examining figure 1, it is apparent that, if the dividend payout percentage and the corporate tax rate are held constant, a decrease in the investor tax rate will increase the incremental tax cost of the corporate form. With ERTA, the maximum individual tax rate decreased from 70 to 50 percent, while the maximum corporate rate remained unchanged. Therefore, if everything else remained constant (dividends, earnings, tax credits, etc.), corporations should have experienced increased incremental tax costs after 1981.[1] This prediction is examined empirically by testing the following null hypothesis for the 1981 ERTA tax rate reductions:

H3: The change in the incremental tax cost of the corporate form after 1981 is less than or equal to zero.

Hypothesis H4

Hypothesis H4 looks at reactions of managers of corporations to increasing incremental tax costs brought about by the 1981 tax rate reductions. One way corporations can reduce their incremental tax cost is to decrease dividend payments. However, emperical evidence indicates that dividend levels remain constant or increase over time. (See Barclay and Smith [1988] for a discussion of dividend policies.) Firms rarely reduce dividends, and when they do, it is usually associated with a decline in earnings. Boards of directors may select optimal dividend levels on the basis of factors uncorrelated with tax effects (e.g., investment opportunities, signaling, free cash flow). For these reasons, any reduction of dividends may be a sub-optimal decision, and thus other alternatives may have been pursued.

Hypothesis H4.1. In discussing the effects of the 1986 Tax Reform Act, Wolfson (1988, 18) states:

> ...it is straightforward to show that the effect of the new tax bill is to encourage increased reliance on corporate borrowing as a means of financing. The result is the same as if the corporation had converted partially to a partnership form of organization for tax purposes.

This result follows from the tax deductibility of interest payments. To the extent the firm substitutes debt holders for equity holders, it can provide a return to its suppliers of capital (debt holders) that is subject to only a single level of taxation, the same as a partnership. The additional interest payments will reduce the incremental tax cost because of their deductibility by the corporation.

[1] For example, under the assumptions of figure 1, if the individual tax rate is decreased from 60 to 50 percent, the tax cost of the corporation per $1 of taxable income is 0.460+0.135=0.595, while the tax cost of the partnership is 0.500 for an incremental tax cost of 0.095.

If competition from partnership forces managers of corporations to increase debt levels (thereby reducing tax costs), this might imply that managers would otherwise select debt levels that did not minimize tax costs, a potential criticism of hypothesis H4.1. However, as Jensen (1986) and Summers (1989) point out, adding debt to a corporation's capital structure imposes costs on managers, and this in turn may cause managers to make capital structure decisions that are not cost-minimizing for the firm. For example, Summers (1989, 302) states, "managers do not always act in the best interests of their shareholders—the 'agency' problem. Running a highly leveraged corporation may be more profitable for shareholders but less pleasant for managers." However, when managers are faced with a greater potential cost (i.e., the loss of their job) due to competition from the partnership form of organization, they may be forced to take actions that increase their own costs but reduce the tax cost of the corporation.

On the basis of this analysis, corporations with increasing incremental tax costs should also be observed increasing their long-term debt levels, which is tested by hypothesis H4.1. Stated in the null form:

H4.1: The correlation between the change in a corporation's incremental tax cost and the change in the corporation's long-term debt after 1981 is less than or equal to zero.

Hypothesis H4.2. Another way corporations could deal with increasing incremental tax costs is to attempt to distribute assets out of the corporation directly to their shareholders (a form of gradual liquidation). Transaction costs prevent the firm from simply liquidating and forming a partnership. However, it may maximize the shareholders' wealth if the firm gradually distributes its assets and enables shareholders to convert their investment into a non-corporate form.[2]

These distributions, which generally take the form of stock repurchases, are known as non-dividend distributions and are usually taxed to shareholders as capital gains. According to an empirical observation by Shoven (1986, 29, 31), "corporations in the United States are now making non-dividend cash payments to shareholders, the sum of which exceeds that of dividends. . . . Clearly there has been a major change in corporate financial behavior which necessitates future research." Although Shoven speculates that tax minimization behavior may be one cause of this behavior, he does not conduct any specific tests of this hypothesis, nor does he relate the behavior to the 1981 tax rate changes.

Another type of non-dividend distribution is cash payments by an acquiring corporation to target shareholders in a merger. Petruzzi (1988, 104) develops a model in which changes in tax laws lead to a change in the relative values of investment opportunities in the corporate and non-corporate sectors. He argues that "a relative tax cut for individuals makes it less likely that capital will flow into corporations and more likely that it will flow out." Petruzzi uses this result as an explanation for why firms make cash payments to target firm shareholders in mergers. The reponse of managers to increasing stock repurchases or increasing merger payments is tested by hypothesis H4.2. Stated in the null form:

---

[2] Note that these non-dividend distributions cannot completely take the place of ordinary dividends. Otherwise, firms could simply stop paying dividends and substitute non-dividend distributions. One reason for this lack of substitutability is the Internal Revenue Code. For example, Section 302 would re-characterize a *pro rata* stock repurchase as an ordinary dividend. Barclay and Smith (1988) argue that opportunistic behavior by managers prevents non-dividend distributions from being perfect substitutes for cash dividends.

> H4.2: The correlation between the change in a corporation's incremental tax cost and the change in the corporation's non-dividend distributions after 1981 is less than or equal to zero.

Hypothesis H4.3. A third possible response to increasing incremental tax costs is a decrease in the dividend payout ratio.[3] As discussed above, one way a corporation could reduce its incremental tax cost is by reducing the level of dividends. However, firms choose a level of dividends for a variety of reasons that may be unrelated to income tax effects. Once a given level of dividends is chosen, the firm may be reluctant to reduce either the per share or total amounts. However, the firm may still achieve a reduction in the incremental tax cost by reducing the payout ratio. For a firm with increasing income, the dividend payout ratio may be lowered without reducing either per share or total dividends simply by keeping the dividend level constant. This response is tested by hypothesis H4.3. Stated in the null form:

> H4.3: The correlation between the change in a corporation's incremental tax cost and the change in the corporation's dividend payout ratio after 1981 is greater than or equal to zero.

Although hypotheses H4.1, H4.2, and H4.3 are all stated separately, a firm is not constrained to only one response. Firms may in fact respond to increasing incremental tax costs by some combination of increasing long-term debt, increasing non-dividend distributions, and decreasing dividend payout ratios. For this reason, both univariate and multivariate tests of hypothesis H4 are conducted.

II. Measurement of Variables and Sample Selection

Measuring Incremental Tax Cost

A measure of the incremental tax cost of the corporate form, labelled the tax wedge here,[4] is required to test hypotheses H1, H3, and H4. The tax wedge is defined as the difference between (1) the net after-tax increase in wealth accruing to the partners of a partnership, and (2) the sum of undistributed net income and after-tax dividends accruing to the shareholders of a corporation, assuming both the corporation and the partnership have the same pre-tax and taxable income.

To derive the tax wedge, let EBT be the net earnings before tax (assumed for simplicity to be equal to cash flow) and TI be the taxable income of an entity. If the entity is a corporation, its undistributed net income (UNI) is then:

$$UNI = EBT - TI(t_c) - D, \tag{1}$$

where t_c is the corporate tax rate and D is the amount paid as a dividend.

---

[3] The dividend payout ratio is defined as the ratio of total dividends to pre-tax book income. The dividend level is defined as the total dollar amount of dividends paid.

[4] The term "tax wedge" is borrowed from the public finance sector of the economics literature, where it refers to the difference between the price received by the seller of a good and the price paid by the buyer of a good when the difference is caused by a tax. In this study, it refers to the difference in tax costs of two alternative organizational forms.

The after-tax increase in wealth accruing to the shareholders ($\Delta WEALTH_s$) is:

$$\Delta WEALTH_s = UNI - \frac{UNI(t_{cg})}{(1+i)^T} + D(1-t_o), \tag{2}$$

where t_o equals the investor tax rate on ordinary income, t_{cg} is the investor tax rate on capital gains, T is the number of years that the investor holds the stock before the capital gain is taxed, and i is the investor discount rate.

Substituting equation (1) into equation (2) and rearranging gives:

$$\Delta WEALTH_s = EBT - TI(t_c) - \frac{UNI(t_{cg})}{(1+i)^T} - D(t_o). \tag{3}$$

If the entity is a partnership,[5] the after-tax increase in wealth accruing to the partners is:

$$\Delta WEALTH_p = EBT - TI(t_o), \tag{4}$$

since all of the income is taxed directly to the individual partners. No present-value terms are necessary in this expression, since a partner could sell the ownership interest at any time (and buy it back) with no additional tax consequences. Also, note that the level of dividends has no effect on the partnership tax.

The tax wedge (W) is defined here as the difference between equations (4) and (3):

$$W = \Delta WEALTH_p - \Delta WEALTH_s = TI(t_c) - (TI - D)(t_o) + \frac{UNI(t_{cg})}{(1+i)^T} \tag{5}$$

In substance, equation (5) says the tax wedge is equal to the actual corporation income tax [$TI(t_c)$], less the investors' tax on undistributed taxable income [$(TI-D)(t_o)$], plus the present value of the capital gains tax due upon a future sale of the shares [$UNI(t_{cg})/(1+i)^T$].[6] Because of the lower tax rates applicable to capital gains and because of the ability of shareholders to defer the capital gains tax indefinitely, it is initially assumed that the present value of the future tax is close to zero and can be ignored. (This is a common assumption in the finance literature. See Miller [1977] and Miller and Scholes [1978] for examples of this assumption.) The effect of this assumption is to understate the tax wedge and to bias the results against rejection of the null form of hypothesis H1.[7]

Two of the terms in equation (5), TI (the corporation's taxable income) and t_o (the investor marginal tax rate on ordinary income[8]), are not directly observable or available to the researcher. The methods used to estimate these terms are discussed below.

---

[5] This analysis also assumes that the marginal tax rates of the partners are equal to those of the shareholders. This would be the case if the corporation could costlessly change its form to a partnership or if the same investors were considering alternative forms of investment.

[6] This calculation ignores any appreciation in the value of the assets owned by the firm, since this appreciation is not subject to income tax until "realized" for tax purposes by a sale or exchange transaction. Any appreciation in asset values would result in a capital gains tax to the shareholder or partner when the stock or partnership interest was sold. Since the effect would be the same for both the corporate and partnership forms, it is ignored in the calculation of the tax wedge.

[7] The sensitivity of the results to this assumption are examined in section III through the use of a second tax wedge measure, called the upper bound tax wedge. In the calculation of the upper bound tax wedge, the assumption is made that all capital gains are taxed currently.

[8] Marginal rather than average investor tax rates are required in the tax wedge calculation, since the calculation assumes that investors would increase their taxable income by their share of the corporation's undistrib-

Table 1
Weighted Average Marginal Tax Rates
for Individuals Receiving Dividend Income

Year	*Tax Rate*
1978	44.29%
1979	45.26%
1980	46.47%
1981	45.89%
1982	38.41%
1983	35.94%
1984	35.19%
1985	34.98%

Note: The tax rates represent the weighted average marginal U.S. tax rates for individual shareholders who received taxable dividends during the calendar years indicated. Rates are calculated on the basis of data published in the Internal Revenue Service publication "Statistics of Income."

Measuring Investor Tax Rates

The first method of estimating investor tax rates is based on information contained in the Internal Revenue Service publication "Statistics of Income." This publication reports dividends (and other gross income items) received by taxpayers within specific levels of taxable income (classified by filing status). Based on the assumption that the composition of investor groups does not differ across firms, a weighted average investor marginal tax rate is calculated and used. For example, if the IRS data indicate that 10 percent of all dividends are received by taxpayers with a 35 percent marginal tax rate, the assumption is made that 10 percent of the equity owners of any particular corporation have that same rate. A key assumption in this calculation of investor tax rates is the absence of clienteles. If this assumption is invalid, the estimated tax rates will not be applicable to all clienteles. For this reason, attempts to minimize the effect of clienteles on the sample are discussed below. Table 1 gives the weighted average marginal tax rates, calculated with the above method, for each year from 1978 through 1985. The effect of the ERTA rate reductions, which became effective in 1982, is apparent from the table. The gradual decrease in rates after 1982 reflects the fact that rate reductions were "phased-in" over several tax years.[9]

The second method of estimating investor tax rates is based on the change in the price of a corporation's common stock on the day the stock begins trading ex-dividend. This method results in a measure that is specific to the firms in the sample, but is subject to more measurement error than the first method. As discussed by Elton and Gruber (1970), if dividend income and capital gains are taxed at different rates, then

---

uted taxable income. Because of progressive tax rates, this hypothetical additional income would be taxed at rates higher than the average rate.

[9] To examine the sensitivity of the results to the assumption that investors are all individuals, the equation for the tax wedge is re-derived by assuming that a fixed percentage (equal to $1-\Lambda$) of all investors are tax-exempt organizations and that these exempt organizations are subject to the tax on unrelated business income at rates equal to the corporate tax rate. Under these assumptions, the resulting tax wedge is equal to Λ times the original tax wedge (i.e., ΛW). Since Λ is a constant, percentage changes in the tax wedge and t-statistics for the significance of the mean tax wedge are the same for both W and ΛW.

the change in the price of a share of stock on the ex-dividend day will be equal to $D(1-t_o/1-t_{cg})$. To reduce heteroscedasticity, ex-dividend day returns rather than price changes are used. Using ex-dividend day returns results in the following relationship:

$$R_t = \frac{D}{P_{t-1}} \frac{(t_o - t_{cg})}{(1 - t_{cg})}, \tag{6}$$

where R_t is the return expected on a share of stock on the ex-dividend day caused solely by differential tax rates[10] and P_{t-1} is the price of the stock on the day before the ex-dividend day.

To solve for t_o in equation (6), it is necessary to make an assumption about the relationship between t_o and t_{cg}. It is assumed that the statutory relationship between ordinary income tax rates and capital gains tax rates ($t_{cg} = .4t_o$) holds for the marginal investor.[11]

Measuring Corporate Taxable Income

The calculation of $(TI - D)t_o$ in equation (5) requires knowledge of the corporation's taxable income. Since this information is not provided in the corporation's financial statements, a new method of estimating taxable income is developed in this study.[12] This new method starts with the current federal and foreign tax payable and derives federal and foreign taxable income by taking into account the statutory U.S. tax rate, the foreign rate differential, the capital gains tax rate differential, and the amount of tax credits. (These items are chosen since they are typically disclosed in firms' financial statement footnotes and account for the majority of the differences between the statutory and effective tax rates.) The major limitation of this method for researchers is that it requires footnote data that are not available on COMPUSTAT. The details of this computation are presented in the appendix.

The Sample Firms

The Tax Wedge Sample of Corporations. The sample corporations used to measure the tax wedge are chosen by first obtaining a listing of all New York Stock Exchange COMPUSTAT corporations in the manufacturing, wholesale, and retail SIC classifications.[13] Corporations without price or dividend data for each year from 1978 through 1985 are deleted, as are corporations that did not pay an annual dividend of at least $0.10 per share every year from 1978 through 1985.[14] The sample corporations are also constrained to have calendar year ends.[15]

---

[10] Since a share of stock is expected to have a positive daily return without considering tax effects (the "normal return" from an asset-pricing model), equation (6) represents the return expected on the ex-dividend day in addition to the normal return (i.e., the "abnormal return"). Abnormal returns are calculated by using the one-factor market model and the value-weighted market index, with parameters estimated over the 250 trading days prior to the month in which the stock goes ex-dividend. Estimates were also made with the equally weighted market index, as well as with market-adjusted returns (i.e., $Ri - Rm$).

[11] However, if the marginal investor is a short-term trader rather than a long-term investor, this assumption is less likely to hold, since short-term traders will be either corporations or tax-exempt organizations. Short-term trading activity is expected to be greatest in high dividend yield stocks (since the positive ex-dividend day return captured by short-term traders is a function of the dividend yield). As discussed in section III, this study finds empirical evidence that, for sample firms with dividend yields above the median, changes in ex-dividend day returns are inconsistent with the 1981 tax rate reductions. This result could be evidence of short-term trading activity in high-yield stocks.

[12] One way to estimate taxable income (used in previous studies) is to start with pre-tax book income and adjust by the change in the deferred tax balance sheet accounts divided by the U.S. statutory tax rate. That method is not used here because it ignores the effect of differential foreign tax rates, ignores the effect of permanent differences, assumes that deferred taxes are provided and reversed at the statutory tax rate, and does not discriminate between ordinary taxable income and capital gains.

The annual dividend yield (based on year-end stock price) is calculated for each firm-year in the sample. The mean yield (across all firm-years) is 0.03785 with a standard deviation of 0.02111. Selecting sample corporations with similar dividend yields helps control for tax clientele effects on shareholder tax rates.[16] To get corporations with dividend yields as similar as possible, firm-years in which the dividend yield was more than one standard deviation above or below the mean are eliminated. This procedure results in a sample of 100 corporations.[17]

From the 100 corporations representing all corporations with eight years of complete data (1978-85) and meeting the dividend yield restriction, three corporations are deleted because of a lack of tax footnote information, and two corporations are deleted because they have such large net operating loss carry-forwards and investment tax credit carry-forwards that calculating tax wedges is not possible. This leaves a final sample of 95 corporations, yielding 760 firm-years of observations. Table 2 reports in summary form how the final sample is arrived at, setting out the number of corporations eliminated at each of the stages discussed above.

Summary statistics showing total assets, total sales, and debt/asset ratios for the sample corporations for the eight-year sample period are reported in table 3. The table reports that measures of both assets and sales increased for the sample corporations from 1978 to 1985, with mean (median) assets increasing by 101 percent (112 percent) and mean (median) sales increasing by 89 percent (70 percent). Because of this growth in firm size over the sample period, taxable income (and thus the tax wedge) is also expected to grow. To eliminate the effect of this increase from the tax wedge measures in the empirical tests, the tax wedge is scaled by total sales.

The Sample of Master Limited Partnerships. To test hypothesis H2, a sample of MLPs is identified by using Moody's "Tax Status of 1987 Dividends," which gives information about tax-free dividends paid by publicly held companies. Since cash distributions made by partnerships are not subject to tax, they are treated as tax-free dividends by Moody's. With this method, 86 MLPs are identified for 1987. Of these, 13 do not have data on COMPUSTAT, seven had no operations for 1987, nine are in industries concerned with investments (e.g., mortgage funds, equipment leasing, real estate leasing), four were in the process of liquidating, and 16 could not be matched

---

[13] Corporations in the extractive industries are not used because of the presence of unique tax rules available only to those industries (e.g., percentage depletion, expensing of intangible drilling or mining costs). For similar reasons, financial institutions and insurance companies are also excluded.

[14] This restriction is necessary in order to observe ex-dividend day price changes, as discussed above. Since share prices generally change only by increments of $0.125, extremely small per-share dividend amounts are not expected to be reflected in price changes.

[15] This is necessary since the equation for the tax wedge assumes that dividends paid will be taxed at the same marginal investor tax rate as partnership income would be, and this assumption will not hold for fiscal year corporations if investor tax rates differ across calendar years.

[16] Clientele theory states that shareholders with high marginal tax rates will tend to invest in stocks with low dividend payouts, while shareholders with low marginal tax rates will tend to invest in stocks with high dividend payouts.

[17] Since the selection of one standard deviation as a cutoff point is arbitrary, the effect of this selection procedure on the results of the hypotheses tests is examined in section III.

Table 2
Sample Selection Statistics for the Tax Wedge Sample of Corporations[a]

Sample Selection Summary	
Initial COMPUSTAT sample[b]	1,267
Less: Firms not on the New York Stock Exchange	458
	809
Less: Firms with missing price or dividend data or firms with dividends less than $0.10	278
	531
Less: Firms with other than a calendar year end	209
	322
Less: Firms with dividend yields more than one standard deviation from the mean[c]	122
	100
Less: Firms with missing footnote information (3) or net operating loss carry-forwards (2)	5
Final Sample Size	95

[a] This sample of corporations is used to estimate tax wedges and in the empirical tests of hypotheses H1, H3, and H4.

[b] Includes only firms from the COMPUSTAT industrial file in the manufacturing, retail, and wholesale industry classifications.

[c] Mean is calculated for 531 firms (before eliminating non-calendar year ends).

with a corporation in that same industry (e.g., the Boston Celtics). This leaves 37 MLPs in the sample, in eight different industries (measured by four-digit SIC codes). Of these 37 MLPs, 18 also had operations in 1986. Prior to 1986, too few MLPs had operations (seven) to allow meaningful statistical tests. Table 4 reports in summary form how the final sample was arrived at, setting out the number of MLPs eliminated at each of the stages discussed above.

The Sample of Matched Corporations. Ideally, the sample of MLPs would be compared with the sample of corporations used for the tax wedge tests of hypothesis H1 to determine cross-sectional differences in transaction costs. However, the sample of MLPs consists of firms with an average size (measured by assets) approximately one-third that of the sample of corporations used in the estimation of the tax wedges. Also, the industry composition of the two samples is different, since the MLPs are primarily in the oil and gas industry, while the corporations are selected by excluding firms in extractive industries. Transaction costs may differ across different sizes of firms, as well as across different industries.

To control for the difference in size and industry between the sample of MLPs and the sample of corporations used for the tax wedge calculations, a comparison sample of 37 corporations (of which 18 are also in 1986) is selected from COMPUSTAT firms

Table 3
Summary Statistics for Sample Corporations Used for Tax Wedge Calculations
(Dollar amounts in Millions)

Year	*Mean (Median) Total Assets*	*Mean (Median) Total Sales*	*Mean (Median) Debt/Assets*
1978	1,132	1,473	0.438
	(549)	(761)	(0.444)
1979	1,329	1,763	0.443
	(696)	(880)	(0.453)
1980	1,507	2,025	0.443
	(786)	(1,000)	(0.443)
1981	1,690	2,249	0.445
	(879)	(1,046)	(0.437)
1982	1,776	2,270	0.447
	(895)	(963)	(0.453)
1983	1,905	2,434	0.444
	(967)	(1,102)	(0.446)
1984	2,014	2,655	0.451
	(1,090)	(1,272)	(0.446)
1985	2,274	2,779	0.472
	(1,165)	(1,295)	(0.463)

Note: This sample of corporations is used to estimate tax wedges, and in the empirical tests of hypotheses H1, H3, and H4.

matched with the MLPs by firm size (as measured by assets) and industry membership (four-digit SIC code). (Note that, because of the size and industry differences discussed above, these corporations are not the same as those in table 2.)

Summary statistics showing total assets, total sales, and debt/asset ratios for the sample of MLPs and the matched corporations for 1986 and 1987 are given in table 5. The table shows that the average total assets and average sales for the MLPs and matched corporations are not significantly different from each other. Table 5 also reports that the MLPs in the sample have significantly less debt (at the 0.10 level) than the matched corporations, as measured by the debt/asset ratio. This is consistent with the fact that MLPs do not receive the tax benefit from deducting interest expense that corporations do. This difference in leverage is controlled for in the empirical tests.

III. Empirical Tests of Hypotheses

Hypothesis H1

To test hypothesis H1, a mean tax wedge (W_i^*) is calculated for each firm i in the sample for the entire eight-year sample period. The hypothesis test is based on whether the mean of this distribution (labelled W^*) is greater than zero. The null form of hypothesis H1 is $W^* \leq 0$. This hypothesis is tested by using a t-test for the significance of the mean, and the result is presented in table 6. Results for two forms of the tax wedge

Table 4
Selection Statistics for the Sample of Master Limited Partnerships

Sample Summary

Firms identified from Moody's Tax Status of 1987 Dividends	86
Less: Firms not on COMPUSTAT	13
Firms with no operations	7
Firms in investment activities	9
Firms with no matching corporation	16
Firms in liquidation	4
Total Sample Firms	37

Note: This sample of master limited partnerships is matched by size and industry with a sample of corporations, and the matched pairs are used in the empirical tests of hypothesis H2.

Table 5
Summary Statistics for Matched Samples of Corporations and MLPs[a]
(Dollars in Thousands)

Mean Values:	*MLPs*	*Matched Corporations*	*t-statistic*[b] (*p-values*)
1986			
Total Assets	717,647	654,209	−1.29 (0.221)
Total Sales	194,817	196,359	0.05 (0.960)
Debt/Assets	0.433	0.657	1.92 (0.076)
Number of Firms	18	18	
Number of Industries	1	1	
1987			
Total Assets	479,705	461,297	−0.50 (0.623)
Total Sales	214,885	175,633	−0.56 (0.556)
Debt/Assets	0.489	0.625	1.86 (0.074)
Number of Firms	37	37	
Number of Industries	8	8	

[a] The sample of master limited partnerships is matched by size and industry with a sample of corporations, and the matched pairs are used in the empirical tests of hypothesis H2.

[b] t-statistics test the null hypothesis that the means of the two samples are equal; p-values are for two-tailed tests.

Table 6
Test of Hypotheses H1 and H3[a]
Mean Tax Wedge/Sales

Period	Percentiles 10%	50%	90%	*Mean* *(Std. Dev.)*	*t-statistic*[b] *(p-value)*
1978–85					
Wedge$_1$[c]	0.0033	0.0102	0.0229	0.0119 (0.0075)	15.48 (0.001)
Wedge$_2$	−0.0057	0.0030	0.0104	0.0030 (0.0062)	3.15 (0.002)
1978–81					
Wedge$_1$	0.0027	0.0092	0.0203	0.0103 (0.0070)	14.37 (0.001)
Wedge$_2$	−0.0129	−0.0044	0.0036	−0.0041 (0.0070)	−3.92 (0.999)
1982–85					
Wedge$_1$	0.0047	0.0110	0.0256	0.0135 (0.0087)	15.20 (0.001)
Wedge$_2$	0.0010	0.0099	0.0192	0.0101 (0.0081)	8.23 (0.001)
Difference					
Wedge$_1$				0.0032	6.56[d] (0.001)
Wedge$_2$				0.0142	10.97[d] (0.001)

[a] Hypothesis H1 (null form) states that the tax wedge is less than or equal to zero; hypothesis H3 (null form) states that the change in the tax wedge after 1981 is less than or equal to zero.

[b] *t*-statistics test the hypothesis that the mean is equal to zero; *p*-values are based on one-tailed tests.

[c] Wedge$_1$ uses estimates of investor tax rates based on IRS Statistics of Income, while those of Wedge$_2$ are based on ex-dividend day price changes.

[d] Here, *t*-statistics test the hypothesis that the difference is equal to zero.

are reported—Wedge$_1$ uses estimates of investor tax rates based on IRS Statistics of Income, and Wedge$_2$ uses estimates of investor tax rates based on an ex-dividend day price changes. Table 6 reports that the mean tax wedge for both Wedge$_1$ and Wedge$_2$ is significantly greater than zero for the sample period; thus, the null form of hypothesis H1 is rejected. During the period 1978–85, firms of the type represented in the sample faced a greater tax cost by choosing the corporate form of organization.

It should be noted that only half of the sample firms are used for the Wedge$_2$ tests (44 firms after eliminating those with missing returns data). This is because of an empirical finding in this study that the change in estimated tax rates after 1981 was significantly different for firms with low and high dividend yields. The low-yield firms exhibited the expected decrease in estimated tax rates after 1981, but the high-yield firms exhibited an increase in estimated tax rates after 1981. Since previous empirical and theoretical research indicates that ex-dividend day returns of high-yield stocks are expected to be affected by short-term trading activity (see, e.g., Karpoff and Walkling 1988), only the returns from the low-yield firms are used.

One unusual result apparent in table 6 is the negative tax wedge for the 1978–81 period when ex-dividend day returns are used to estimate tax rates (Wedge_2). Dividend clientele theory suggests that investors in low dividend yield firms are expected to be high marginal tax rate investors. Consistent with this expectation, the estimated tax rates using ex-dividend day returns from the low-yield firms are high (61.5 percent in 1978–81 and 42.5 percent in 1982–85). After examining equation (5), it is apparent that a combination of low dividends and high investor tax rates tends to reduce the tax wedge. One argument against corporations' having a greater tax cost than partnerships is that investors form dividend clienteles to minimize this cost, and, at least for some investors (high tax rate investors and low dividend yield stocks), this appears to have been possible prior to 1981. However, it is apparent from table 6 that after 1981, even with dividend clienteles, the corporate form had a tax cost greater than the partnership form.

Hypothesis H2

In conformance with Karpoff and Rice (1989), transaction costs for hypothesis H2 are measured with accounting-based performance measures. The four measures used are (1) return on total assets (EBIT/A), equal to net income before interest expense and taxes, divided by the book value of total assets, (2) return on sales (EBIT/S), equal to net income before interest expense and taxes, divided by total sales, (3) operating income to sales (OPI/S), equal to operating income divided by total sales, and (4) gross return on sales (GP/S), equal to gross profit divided by total sales. In addition to these four performance measures, a fifth measure equal to total selling, general, and administrative expense divided by total sales (SGA/S), is used to test for the presence of transaction costs in selling, general, and administrative expense.

The performance measures are calculated for each of the MLPs and matched corporations in the sample. On the basis of the alternative form of hypothesis H2 (partnerships experience greater transaction costs than corporations), the performance measures of the partnerships are predicted to be less than those of the matched corporations. Also, partnerships are predicted to have greater selling, general, and administrative expenses than the matched corporations. Differences in mean values are tested by using a matched pair *t*-test.

Table 7 reports the results of testing hypothesis H2. For 1986, all five of the mean differences are in the predicted direction. Of the performance measures, EBIT/A is significant at the 0.10 level, while EBIT/S, OPI/S, and GP/S are significant at the 0.05 level. The difference in SGA/S is not significant. For 1987, the first two performance measures, EBIT/A and EBIT/S, are in the predicted direction and significant at the 0.05 level. The other two performance measures, OPI/S and GP/S, as well as SGA/S, are not in the predicted direction.

Taken together, these results indicate that the MLPs have lower accounting-based return on assets and return on sales than the matched corporations. The results are consistent with greater transaction costs for MLPs than for corporations. In 1986, these transaction costs are reflected in cost of goods sold and result in a lower gross profit percentage. For 1987, the gross profit and operating income amounts for the two types

Table 7
Test of Hypothesis H2[a]
Comparison of Accounting-Based Performance Measures for Samples of MLPs and Matched Corporations

Performance Measure[b]	Mean [Median] Corporation	Mean [Median] MLP	Mean of Differences (Corp. − MLP)	Expected Sign	t-statistic	p-value[c]
1986 Data						
EBIT/A	−0.047 [0.043]	−0.196 [−0.215]	0.149	(+)	1.49	0.076
EBIT/S	−0.117 [0.141]	−0.746 [−0.776]	0.629	(+)	2.16	0.024
OPI/S	0.285 [0.403]	−0.185 [0.064]	0.470	(+)	1.85	0.035
GP/S	0.444 [0.556]	−0.002 [0.225]	0.446	(+)	1.89	0.038
SGA/S	0.179 [0.167]	0.182 [0.192]	−0.003	(−)	−0.07	0.472
1987 Data						
EBIT/A	0.083 [0.059]	0.039 [0.051]	0.044	(+)	1.80	0.041
EBIT/S	0.191 [0.176]	0.036 [0.155]	0.155	(+)	2.43	0.011
OPI/S	0.336 [0.373]	0.367 [0.390]	−0.031	(+)	−0.73	0.770
GP/S	0.495 [0.531]	0.508 [0.571]	−0.013	(+)	−0.26	0.598
SGA/S	0.178 [0.187]	0.153 [0.148]	0.025	(−)	1.11	0.860

[a] Hypothesis H2 (null form) states that the incremental transaction costs of the partnership form over the corporate form are less than or equal to zero.

[b] EBIT/A is equal to net income before interest and taxes divided by total assets; EBIT/S is equal to net income before interest and taxes divided by total sales; OPI/S is equal to operating income divided by total sales; GP/S is equal to gross profit divided by total sales; SGA/S is equal to selling, general, and administrative expense divided by total sales.

[c] p-values are based on one-tailed tests.

of firms are not significantly different, but the net income before taxes and interest is lower for MLPs than for corporations. There is no evidence in either 1986 or 1987 that MLPs incur greater selling, general, and administrative expenses than corporations.[18]

MLPs are a new form of business entity, and new businesses may be run less efficiently than older ones simply because they are new. In additon, newer businesses may have higher book values (for comparable assets) than older businesses simply because

---

[18] Additional tests are performed to determine if these results are caused by differences in accounting methods. These tests indicate that differences in accounting methods do not explain the results reported in table 7.

prices are rising. However, although MLPs are a new organizational form, they are not new businesses, since the sample MLPs (and MLPs in general) are formed by combining several existing non-traded limited partnerships into one organization (a "roll up") or by spinning off a segment of an existing corporation (a "roll out"). In both cases, the actual day-to-day operations of the business are not changed. To control for MLPs having more assets on the books, the pairs of firms in the sample are matched on total assets. Also, the sample MLPs have a higher (though not significantly so) sales/assets ratio than the sample corporations for both 1986 and 1987. Therefore, differences in performance do not appear to be caused by higher asset book values for the MLPs.

Two sets of biases are expected to influence the results for hypothesis H2, and these are expected to be in opposite directions. First, the fact that the MLPs in the sample "self-selected" the partnership (rather than the corporate) form means that only those firms for which the expected transaction costs are low will choose to be partnerships. This will bias the results of the hypothesis tests against rejecting the null hypothesis. The second possible bias occurs because the MLPs in the sample may have chosen this form in anticipation of the greater tax costs of the corporate form arising from the Tax Reform Act of 1986. If so, then the transaction costs measured in 1986 and 1987 may be greater than those that would have existed prior to 1986. These greater transaction costs will bias the results in favor of rejecting the null hypothesis.[19] Since these two potential biases tend to offset one another, it is impossible to assess the effect of each one individually on the test results.

Hypothesis H3

To test hypothesis H3, the average tax wedge (W_i^*) is calculated for each sample corporation i over two sub-periods, 1978–81 and 1982–85, which gives two averages for each firm, $W_i^{*\,1978\text{-}81}$ and $W_i^{*\,1982\text{-}85}$. Next, the difference between these two averages (ΔW_i) is calculated for each sample corporation i by subtracting $W_i^{*\,1978\text{-}81}$ from $W_i^{*\,1982\text{-}85}$. Finally, the mean of ΔW_i is calculated (ΔW^*), and the null hypothesis is $\Delta W^* \leq 0$. This hypothesis is tested by a one-sample t-test.

The tests of hypothesis H3 are reported at the bottom of table 6. The mean of the differences in the average tax wedges is significantly greater than zero for both forms of the tax wedge. This implies that, on average, the sample firms experienced an increase in the tax wedge (as a percentage of total sales) after 1981. This finding is consistent with the alternative form of hypothesis H3 that the 1981 tax rate reductions resulted in an increase in the tax cost of the corporate form. Also, it is important to note that those low-yield firms with high marginal tax rate investors (Wedge_2) had their tax cost change from being less than the partnership form to being greater than the partnership form after 1981.

Hypothesis H4

In the univariate tests of hypothesis H4, a Pearson correlation coefficient is used to test the significance of the correlations between the change in the tax wedge and changes in long-term debt, non-dividend distributions (both stock repurchases and merger payments), and dividend payout ratios. Two measures of the change in the tax wedge are used in these tests.[20] Both of the measures are based on the average of the tax wedge divided by sales for the sub-periods 1978–81 and 1982–85. The first measure is the percentage change between the two sub-periods.[21] The second measure is the average value of the post-1981 tax wedge/sales. The tax wedges used to test this hypothesis cannot be the same as those used to test hypotheses H1 and H3, since the original tax wedges are all calculated with the actual data for each year in the sample period. If firms actually respond to an increase in tax wedges (e.g., by increasing long-term debt), the tax wedges originally calculated will include the effects of the responses themselves. The effect of the response on the tax wedge is to reduce the increase (in the tax wedge) caused by the tax rate change. This will reduce the correlation between the responses and the increase in the observed tax wedges. To remove the effect of the responses from the tax wedge measures, an "as if" tax wedge is calculated for each firm in the sample for the years 1982–85. In general, the "as if" tax wedge is calculated by assuming that the response variable being investigated (e.g., long-term debt) remained at its average 1978–81 level after 1981.

---

[19] It may be more accurate to state that the level of transaction costs in the sample partnerships after 1985 may not be comparable with the level that would have existed in a sample of partnerships in 1984 or 1985, and therefore the ability to make inferences about the level of transaction costs prior to 1986 may be impaired.

Table 8 reports the results of testing hypothesis H4.1, with long-term debt as the response variable.[22] The top panel of the table reports the results when long-term debt is divided by the book value of total assets, while the bottom panel reports the results when long-term debt is divided by the sum of the book value of debt plus the market value of equity. The results for both measures of long-term debt are similar. The correlation coefficients between percentage changes in long-term debt and percentage changes in the tax wedge are positive and significant at the 0.05 level. Also, the correlation coefficients between percentage changes in long-term debt and the average level of the tax wedge for 1982–85 are positive, and significant at the 0.05 level. These results are consistent with the prediction of hypothesis H4.1 that firms facing increasing tax costs respond by increasing long-term debt.

The results of testing hypothesis H4.2, with non-dividend distributions as the response variable,[23] are reported in table 9. The top panel reports the results for total non-dividend distributions, the middle panel reports the results for merger payments only, and the bottom panel reports the results for stock repurchases only. The correla-

---

[20] Two measures are used since there are two ways in which changes in the tax wedge may relate to firm responses: firms may respond either to the percentage increase in the tax wedge or to the magnitude of the tax wedge that results after the increase. Under the first view, firms continuously attempt to keep their tax wedge at an optimal percentage level through their responses. A small percentage increase in the tax wedge will result in a small response, and a large increase in a large response. Under the second view, firms attempt to reduce their tax wedge only if it increases above some critical absolute level, such as the incremental transaction cost of the partnership form. Any increase in the tax wedge that leaves the resulting tax wedge below this critical level will not result in a response from the firm. However, if even a small increase in the tax wedge causes it to increase above the critical level, a large response by the firm will result.

[21] The results reported are based on "percentage point changes." For example, an increase in the average tax wedge/sales from 0.10 to 0.20 is considered to be a 10 percent increase, not a 100 percent increase. The latter measure (called the "relative percentage change") is also used, but the results are not reported. The results using relative percentage changes are generally not as significant as the results using percentage point changes.

[22] The number of firms in the sample for this test is decreased to 92 for Wedge_1. Two firms are eliminated because missing COMPUSTAT data prevented calculating changes in long-term debt. One firm is eliminated because its percentage change in long-term debt is an extreme outlier, lying 25 standard deviations from the mean. The next largest observation is four standard deviations.

[23] For this test, non-dividend distributions are measured as a percentage of regular dividends. This allows the calculation of a percentage change when the non-dividend distributions for 1978–81 are zero.

Table 8
Test of Hypothesis H4.1
Correlation Between Changes in the "as if" Tax Wedge/Sales and Percentage Changes in Long-Term Debt
(Predicted Correlation is Positive)

Panel A. Percentage Change in the Ratio of Long-term Debt to Book Value of Total Assets:

Tax Wedge Measures	*Correlation Coefficient*	*t-statistic*	*p-value*
Percentage change in tax wedge (1978–81 to 1982–85)			
$Wedge_1$ / Sales (n=92)	0.206	2.00	(0.025)
$Wedge_2$ / Sales (n=44)	0.299	2.03	(0.026)
Levels of tax wedge (1982–85)			
$Wedge_1$ / Sales (n=92)	0.254	2.49	(0.008)
$Wedge_2$ / Sales (n=44)	0.389	2.74	(0.005)

Panel B. Percentage Change in the Ratio of Long-term Debt to the Sum of the Book Value of Debt Plus the Market Value of Equity:

Tax Wedge Measures	*Correlation Coefficient*	*t-statistic*	*p-value*
Percentage change in tax wedge (1978–81 to 1982–85)			
$Wedge_1$ / Sales (n=92)	0.308	3.08	(0.001)
$Wedge_2$ / Sales (n=44)	0.317	2.16	(0.021)
Levels of tax wedge (1982–85)			
$Wedge_1$ / Sales (n=92)	0.311	3.10	(0.001)
$Wedge_2$ / Sales (n=44)	0.349	2.41	(0.011)

Note: t-statistics test the hypothesis that the correlation coefficient is equal to zero; p-values are based on one-tailed tests. $Wedge_1$ uses estimates of investor tax rates based on IRS Statistics of Income; $Wedge_2$ uses estimates based on ex-dividend day price changes.

tion coefficients between changes in the tax wedge and changes in all three measures of non-dividend distributions are positive and significant at the 0.05 level, with two exceptions.[24] These results are consistent with the prediction that firms facing increasing tax wedges respond by increasing their non-dividend distributions, and that both merger payments and stock repurchases are used for this purpose. The results for merger payments are consistent with the analytical work of Petruzzi (1988), who predicts that an increase in the relative tax cost of corporations compared with partnerships will result in an increase in merger payments.

---

[24] The exceptions are the percentage changes for $Wedge_2$ when merger payments are used ($p=0.110$) and the level of $Wedge_2$ when stock repurchases are used ($p=0.065$).

Table 9
Test of Hypothesis H4.2
Correlation Between Changes in the "as if" Tax Wedge/Sales and Percentage Changes in Non-Dividend Distributions/Dividends
(Predicted Correlation is Positive)

Panel A. Percentage Change in the Ratio of Non-dividend Distributions/Dividends:

Tax Wedge Measures	*Correlation Coefficient*	*t-statistic*	*p-value*
Percentage change in tax wedge (1978–81 to 1982–85)			
Wedge$_1$ / Sales ($n=89$)	0.738	10.19	(0.001)
Wedge$_2$ / Sales ($n=44$)	0.361	2.51	(0.008)
Levels of tax wedge (1982–85)			
Wedge$_1$ / Sales ($n=89$)	0.601	7.14	(0.001)
Wedge$_2$ / Sales ($n=44$)	0.678	5.90	(0.001)

Panel B. Percentage Change in the Ratio of Merger Payments/Dividends:

Tax Wedge Measures	*Correlation Coefficient*	*t-statistic*	*p-value*
Percentage change in tax wedge (1978–81 to 1982–85)			
Wedge$_1$ / Sales ($n=89$)	0.636	7.69	(0.001)
Wedge$_2$ / Sales ($n=43$)	0.191	1.24	(0.110)
Levels of tax wedge (1982–85)			
Wedge$_1$ / Sales ($n=89$)	0.532	5.86	(0.001)
Wedge$_2$ / Sales ($n=43$)	0.590	4.68	(0.001)

Panel C. Percentage Change in the Ratio of Stock Repurchases/Dividends:

Tax Wedge Measures	*Correlation Coefficient*	*t-statistic*	*p-value*
Percentage change in tax wedge (1978–81 to 1982–85)			
Wedge$_1$ / Sales ($n=89$)	0.318	3.13	(0.001)
Wedge$_2$ / Sales ($n=43$)	0.405	2.84	(0.004)
Levels of tax wedge (1982–85)			
Wedge$_1$ / Sales ($n=89$)	0.246	2.36	(0.012)
Wedge$_2$ / Sales ($n=43$)	0.239	1.58	(0.065)

Note: t-statistics test the hypothesis that the correlation coefficient is equal to zero; p-values are based on one-tailed tests. Wedge$_1$ uses estimates of investor tax rates based on IRS Statistics of Income; Wedge$_2$ uses estimates based on ex-dividend day price changes.

Table 10
Test of Hypothesis H4.3
Correlation Between Changes in the "as if" Tax Wedge/Sales and Percentage Changes in the Dividend Payout Ratio
(Predicted Correlation is Negative)

Tax Wedge Measures	*Correlation Coefficient*	*t-statistic*	*p-value*
Percentage change in tax wedge (1978–81 to 1982–85)			
$Wedge_1$ / Sales (n=93)	−0.629	−7.71	(0.001)
$Wedge_2$ / Sales (n=43)	−0.436	−3.10	(0.002)
Levels of tax wedge (1982–85)			
$Wedge_1$ / Sales (n=93)	−0.401	−4.17	(0.001)
$Wedge_2$ / Sales (n=43)	−0.400	−2.80	(0.004)

Note: t-statistics test the hypothesis that the correlation coefficient is equal to zero; p-values are based on one-tailed tests. $Wedge_1$ uses estimates of investor tax rates based on IRS Statistics of Income; $Wedge_2$ uses estimates based on ex-dividend day price changes.

The results of testing hypothesis H4.3 with the dividend payout ratio as the response variable, are presented in table 10.[25] For this hypothesis, the predicted sign of the correlation coefficient is negative. The table reports that all of the correlation coefficients are negative and significant at the 0.05 level. These results are consistent with the prediction that firms respond to increasing tax costs by decreasing the dividend payout ratio.

By examining equation (5), it is apparent that increasing taxable income will cause an increase in the tax wedge (provided the corporate tax rate is greater than the investor tax rate). Also, if a firm has increasing income but is reluctant to increase dividends, the dividend payout ratio will decrease. Therefore, a negative relationship between changes in the tax wedge and changes in the dividend payout ratio could be caused by firms experiencing increased income over the sample period. However, the tax wedge measure used in the empirical tests is divided by sales. An increase in income proportional to an increase in sales will cause the tax wedge/sales to decrease, since the numerator increases less than the denominator (because dividends are constant). Therefore, scaling the tax wedge by sales mitigates any induced negative correlation between the tax wedge and the dividend payout ratio caused by increases in net income.

All of the above tests of hypothesis H4 are univariate tests. However, a firm facing

---

[25] The hypothesis predicts firms will *decrease* dividend payout ratios in response to the 1981 tax rate changes. However, an examination of the data indicates that almost all of the sample firms *increased* their dividend payout ratios over this time period. Therefore, an assumption in the "as if" calculation that all firms kept their dividend payout ratios at the 1978–81 level would be inconsistent with actual economic events. To eliminate the effect of the underlying overall increase in dividend payout ratios and to allow for the examination of the differences in increases across firms, the assumption is made that all firms increased their dividend payout ratios by the average increase. Therefore, a negative correlation indicates a smaller increase (relative to the sample mean), rather than a decrease, in the dividend payout ratio. The sample size for this test is 93 for $Wedge_1$ after elimination of two firms because of missing COMPUSTAT data.

Table 11
Multivariate Reverse Regression of the "as if" Tax Wedge on the Three Response Variables

OLS Reverse Regression: $WEDGE = a + b_1 DEBT + b_2 NDD + b_3 DPR + e$

	a	b_1	b_2	b_3	Adj. R^2
Predicted sign of coefficient		(+)	(+)	(−)	
WEDGE = Percentage change in $Wedge_1$ / Sales ($n = 95$)					
Estimated coefficient:	0.010	0.020	0.005	−0.032	0.534
t-statistic (F-statistic)	6.96	0.95	6.11	−4.69	(36.83)
(p-value)	(0.001)	(0.172)	(0.001)	(0.001)	(0.001)
WEDGE = Percentage change in $Wedge_2$ / Sales ($n = 44$)					
Estimated coefficient:	0.023	0.036	0.006	−0.028	0.338
t-statistic (F-statistic)	8.77	0.81	4.35	−2.37	(8.33)
(p-value)	(0.001)	(0.208)	(0.001)	(0.008)	(0.001)
WEDGE = Level of $Wedge_1$ / Sales for 1982–85 ($n = 95$)					
Estimated coefficient:	0.019	0.043	0.005	−0.027	0.356
t-statistic (F-statistic)	10.91	1.60	5.71	−3.94	(18.35)
(p-value)	(0.001)	(0.056)	(0.001)	(0.001)	(0.001)
WEDGE = Level of $Wedge_2$ / Sales for 1982–85 ($n = 44$)					
Estimated coefficient:	0.016	0.054	0.005	−0.017	0.381
t-statistic (F-statistic)	7.53	1.40	4.56	−2.41	(9.81)
(p-value)	(0.001)	(0.082)	(0.001)	(0.008)	(0.001)

Note: *DEBT* = percentage change in the average ratio of long-term debt to book value of total assets from 1978–81 to 1982–85; *NDD* = percentage change in the average non-dividend distributions from 1978–81 to 1982–85; *DPR* = percentage change in the median dividend payout ratio from 1978–81 to 1982–85. $Wedge_1$ uses estimates of investor tax rates based on IRS Statistics of Income; $Wedge_2$ uses estimates based on ex-dividend day price changes.

increasing tax wedges is not constrained to only one response, and may in fact use a combination of responses. To conduct a joint test for all three responses, the following multivariate test is used. The change in the "as if" tax wedge is used as the dependent variable in an OLS regression. The three independent variables in the regression are the response variables used in the univariate tests: the percentage changes in long-term debt, non-dividend distributions, and the dividend payout ratio. The form of the regression is:

$$WEDGE = a + b_1 DEBT + b_2 NDD + b_3 DPR + e,$$

where *DEBT* is equal to the percentage change in long-term debt/book value of assets, *NDD* is equal to the percentage change in non-dividend distributions/dividends, and *DPR* is equal to the percentage change in the dividend payout ratio.[26]

---

[26] In this form the regression constitutes a "reverse regression," as discussed by Beaver et al. (1987) and Collins and Kothari (1989), since from a causal standpoint the change in the tax wedge (the dependent variable) is assumed to cause the responses (the independent variables). In a reverse regression, the estimates of the slope coefficients represent the upper bound of the inverse of the actual slope coefficients from a forward regression. The sign and significance of the slope coefficient estimates are used to test the predictions of hypothesis H4.

The OLS estimates of the regression coefficients are reported in table 11. The regressions have all been estimated with the White (1980) heteroscedasticity constant covariance matrix to correct for the presence of heteroscedasticity.[27] The *f*-statistics for the regressions are significant, and the coefficients for non-dividend distributions and dividend payout ratios are in the predicted direction and also significant at the 0.05 level. However, the coefficients for long-term debt are significant at the 0.10 level only when levels of the tax wedge are used, and are not significant when percentage changes in the tax wedge are used. Thus long-term debt loses much of its explanatory power when combined with the other two response variables.

Sensitivity Analysis

The sensitivity of the results is tested in three different ways. First, hypotheses are re-tested by using the "upper bound" form of the tax wedge, which is based on the assumption that stockholders are taxed on capital gains in the period in which the gains accrue. The effect of this assumption is to increase the tax wedge in each year. The results of hypotheses tests using the Wedge$_1$ measure are similar to those reported above, except for hypothesis H3, which predicts that tax wedges will increase after 1981. The null hypothesis cannot be rejected with the upper bound Wedge$_1$. However, many individual firms experience increases in their tax wedge over this period, and therefore cross-sectional tests based on changes in the tax wedge are still performed.[28] Also, the actual tax wedge (rather than the "as if" tax wedge) is used for hypothesis H3, which weakens the power of the test. The results for Wedge$_2$ are consistent with the predictions of hypotheses H1 and H3. However, the results using Wedge$_2$ are not generally consistent with the predictions of hypothesis H4.

The second way of testing the sensitivity of the results is to re-test the hypotheses by using a smaller sample of firms with dividend yields closer to the mean dividend yield for all firms. Finding different results may imply that tax clienteles are present in the larger sample. The results for the smaller sample are consistent with and similar to those obtained for the full sample. The major differences are lack of significant results for stock repurchases and greater adjusted R^2 amounts in the multivariate regressions. These results indicate that changing the sample selection dividend yield restriction does not affect the conclusions.

The sensitivity of the parametric test results is verified through the use of computer-intensive randomization procedures described by Noreen (1989). Randomization probabilities are compared with the *p*-values generated by the standard parametric test statistics. If conventional significance levels of 0.10, 0.05, and 0.01 are applied to the results, in no instance is the null hypothesis rejected when using parametric tests but not rejected when using randomization tests. These results imply that the parametric tests are robust to possible violations in distributional assumptions.

IV. Conclusion

This study addresses two research questions dealing with the relationship between income tax laws and organizational form for large, publicly traded business organizations: (1) the incremental costs (including tax costs) associated with different organizational forms, and (2) the actions that managers of existing corporations take in response to an external event (a tax law change) that causes the corporate form of organization to become relatively more costly than an alternative form.

---

[27] This procedure affects the estimates of the standard errors, and hence the *t*-statistics and *p*-values. However, the coefficient estimates are unchanged. The procedure is necessary, since the Breusch-Pagan test rejects the null hypothesis of homoscedasticity at the 0.01 level in each of the regression equations.

[28] The mean Wedge$_1$ in absolute terms does increase after 1981 for the upper bound tax wedge. It is only the tax wedge as a percentage of total sales that does not increase.

The first hypothesis deals with the relative tax cost of the corporate and partnership forms of organization. A tax wedge approach is used to measure the incremental tax cost for a sample of 95 New York Stock Exchange corporations in the manufacturing, wholesale, and retail industries. The results of testing hypothesis H1 reject the null hypothesis for both forms of the tax wedge over the period 1978–85. This means that, for the sample firms, the tax cost of the corporate form was greater than the tax cost that would have been incurred if those same firms had been organized as partnerships.

If the corporate form has a greater tax cost than the partnership form, the corporate form could survive as an alternative organizational form only if there were offsetting costs for the partnership form. The alternative form of hypothesis H2 predicts that the transaction costs of the partnership form are greater than for the corporate form. The hypothesis was tested by comparing accounting-based measures of firm performance between two samples of master limited partnerships and corporations, matched by firm size and industry membership, for 1986 and 1987. In both years, the return on sales and the return on total assets is significantly greater for the corporations than for the partnerships. However, selling, general, and administrative expenses are not greater for partnerships than for corporations.

Hypothesis H3 investigates the effect of a change in investor tax rates on the relative tax costs of corporations and partnerships by examining the effect of the 1981 Economic Recovery Tax Act on the tax wedges of the sample corporations. The alternative form of the hypothesis predicts that corporations experienced increasing incremental tax costs after 1981, and the results of the hypothesis tests are consistent with this prediction. Under hypothesis H4, managers are predicted to take at least three possible actions (responses) to reduce the tax cost of the corporate form. The results of univariate and multivariate tests are consistent with these predictions.

Taken together, these results make a significant contribution to accounting research. They demonstrate a relationship between changes in relative tax costs of alternative (competing) organizational forms and changes in fundamental elements of capital structure, such as debt levels, non-dividend distributions, and dividend payout ratios. One implication of these results is that the relative tax cost of a potential alternative organizational form is likely to affect the capital structure decisions of an existing corporation. A second implication is that a change in tax rates that affects the relative tax costs of competing organizational forms is likely to cause capital structure decisions to change. Previous research on the relationship between taxes and capital structure has not considered the effect of the tax cost of alternative organizational forms. In addition, the results provide empirical support for predictions by Scholes and Wolfson (1986, 1989) (increasing long-term debt) and Petruzzi (1988) (increasing non-dividend distributions) regarding responses of corporations to increases in relative tax costs. The results reported in Guenther (1991) indicate that this relationship between changes in the relative tax cost of alternative organizational forms and changes in corporate capital structure is also observed for the tax rate changes associated with the 1986 Tax Reform Act.

Appendix

A Method for Estimating Corporate Taxable Income with Financial Statement Footnote Information

Definition of Variables Used:

CG = Portion of U.S. taxable income that consists of capital gains.

REC_{CG} = Capital gains reconciling item (expressed as a rate) in the income tax footnote reconciliation of the effective tax rate with the statutory tax rate; usually a negative number. (The statutory tax rates used reflect 1982–85 rates.)

$PTBI$ = Pre-tax book income.

ORD = Portion of U.S. taxable income that consists of ordinary (non-capital gains) income.

$CREDITS$ = Total of all income tax credits.

$CURRENT_{us}$ = U.S. federal income tax expense currently payable.

$RATE_F$ = Effective foreign tax rate.

$BOOK_F$ = Foreign pre-tax book income.

REC_F = Footnote reconciling item for the effect of the difference between the foreign tax rates and the statutory U.S. tax rate.

$CURRENT_F$ = Foreign current tax expense.

DEF_F = Foreign deferred tax expense.

TI_F = Foreign taxable income.

First, estimate the U.S. taxable income from capital gains:

$$CG = \frac{-REC_{CG} \times PTBI}{0.18}. \tag{A1}$$

Second, estimate the U.S. ordinary taxable income:

$$ORD = \frac{CURRENT_{us} + CREDITS - 0.28CG}{0.46}. \tag{A2}$$

Third, estimate the foreign taxable income. This requires several steps.

Step 1:

$$-REC_F = \frac{(0.46 - RATE_F)BOOK_F}{PTBI}. \tag{A3}$$

Rearrange equation (A3):

$$BOOK_F = \frac{REC_F \times PTBI}{(RATE_F - 0.46)}. \tag{A4}$$

Step 2:

$$BOOK_F \times RATE_F = CURRENT_F + DEF_F. \tag{A5}$$

Rearrange equation (A5):

$$BOOK_F = \frac{CURRENT_F + DEF_F}{RATE_F}. \tag{A6}$$

Substitute equation (A6) into equation (A4) and rearrange terms:

$$\frac{CURRENT_F + DEF_F}{RATE_F} = \frac{REC_F \times PTBI}{(RATE_F - 0.46)}; \tag{A7}$$

$$(RATE_F - 0.46)(CURRENT_F + DEF_F) = RATE_F(REC_F \times PTBI); \tag{A8}$$

$$RATE_F(CURRENT_F + DEF_F - (REC_F \times PTBI)) = 0.46(CURRENT_F + DEF_F); \tag{A9}$$

$$RATE_F = 0.46\frac{CURRENT_F + DEF_F}{CURRENT_F + DEF_F - (REC_F \times PTBI)}. \tag{A10}$$

Step 3:

$$TI_F \times RATE_F = CURRENT_F. \tag{A11}$$

Rearrange equation (A11):

$$TI_F = \frac{CURRENT_F}{RATE_F}. \tag{A12}$$

With the above method, the total taxable income for a corporation is the sum of U.S. ordinary income (ORD), U.S. capital gains (CG), and foreign taxable income (TI_F).

References

Arrow, K. J. 1969. *The Organization of Economic Activity: Issues Pertinent to the Choice of Market Versus Nonmarket Allocation,* vol. 1. U.S. Joint Economic Committee, 91st Cong. 1st sess. Washington, D.C.: U.S. Government Printing Office.

Barclay, M. J., and C. W. Smith. 1988. Corporate payout policy: Cash dividends versus open-market repurchases. *Journal of Financial Economics* 22 (October): 61–78.

Beaver, W., R. Lambert, and S. Ryan. 1987. The information content of security prices: A second look. *Journal of Accounting and Economics* 9 (July): 139–57.

Bradley, M., G. Jarrell, and E. Kim. 1984. On the existence of an optimal capital structure: Theory and evidence. *Journal of Finance* 39 (July): 857–77.

Collins, D., and S. P. Kothari. 1989. An analysis of intertemporal and cross-sectional determinants of earnings response coefficients. *Journal of Accounting and Economics* 11 (July): 144–81.

DeAngelo, H., and R. Masulis. 1980. Optimal capital structure under corporate and personal taxation. *Journal of Financial Economics* 8 (March): 3–29.

Elton, E., and M. Gruber. 1970. Marginal stockholder tax rates and the clientele effect. *Review of Economics and Statistics* 52 (February): 68–74.

Fama, E. F., and M. C. Jensen. 1983a. Separation of ownership and control. *Journal of Law and Economics* 26 (June): 301–25.

———, and ———. 1983b. Agency problems and residual claims. *Journal of Law and Economics* 26 (June): 327–49.

Guenther, D. A. 1991. Tax costs, organizational form, and firm capital structure: Evidence from the 1986 tax reform act. Working paper, University of Connecticut, Storrs.

Jensen, M. C., 1986. Agency costs of free cash flow, corporate finance, and takeovers. *American Economic Review* 76 (May): 323–29.

———, and W. H. Meckling. 1976. Theory of the firm: Managerial behavior, agency costs and ownership structure. *Journal of Financial Economics* 3 (October): 305–60.

Karpoff, J. M., and E. M. Rice. 1989. Organizational form, share transferability, and firm performance: Evidence from the ANCSA corporations. *Journal of Financial Economics* 24 (September): 64–106.

———, and R. A. Walkling. 1988. Short-term trading around ex-dividend days: Additional evidence. *Journal of Financial Economics* 21 (September): 291–98.

MacKie-Mason, J. K. 1990. Do taxes affect corporate financing decisions? *Journal of Finance* 45 (December): 1471–93.

Miller, M. H. 1977. Debt and taxes. *Journal of Finance* 32 (May): 261–75.

———, and M. S. Scholes. 1978. Dividends and taxes. *Journal of Financial Economics* 6 (December): 333–64.

Noreen, E. 1989. *Computer-Intensive Methods for Testing Hypotheses: An Introduction.* New York: John Wiley & Sons.

Petruzzi, C. R. 1988. Mergers and the double taxation of corporate income. *Journal of Accounting and Public Policy* 7 (Summer): 97–111.

Scholes, M. S., and M. A. Wolfson. 1986. Taxes and organization theory. Paper presented at the Yale Conference on Economics of Organization.

———, and ———. 1989. Issues in the theory of optimal capital structure. In *Frontiers of Modern Financial Theory,* edited by S. Bhattacharya and G. Constantinides, 49–74. Totowa, NJ: Rowman & Littlefield.

Shoven, J. B. 1986. The tax consequences of share repurchases and other non-dividend cash payments to equity owners. Working paper, Stanford University, Stanford, CA.

Summers, L. H. 1989. Discussion of the impact of the changing tax environment on investments and productivity: Financial structure and the corporation income tax. *Journal of Accounting, Auditing & Finance* 4 (Summer): 302–4.

Titman, S., and R. Wessels. 1988. The determinants of capital structure choice. *Journal of Finance* 43 (March): 1–19.

White, H. 1980. A heteroscedasticity-consistent covariance matrix estimator and a direct test for heteroscedasticity. *Econometrica* 48 (May): 817–38.

Williamson, O. E. 1981. The modern corporation: Origins, evolution, attributes. *Journal of Economic Literature* 19 (December): 1537–67.

Wolfson, M. 1985. Empirical evidence of incentive problems and their mitigation in oil and gas tax shelter programs. In *Principals and Agents: The Structure of Business,* edited by J. Pratt and R. Zeckhauser, 101–25. Boston: Harvard Business School Press.

———. 1988. Taxation, information, and economic organization. *Journal of the American Taxation Association* 9 (Spring): 13–18.

THE ACCOUNTING REVIEW
Vol. 67, No. 1
January 1992
pp. 46–68

Working-Paper Order Effects and Auditors' Going-Concern Decisions

David N. Ricchiute
University of Notre Dame

SYNOPSIS AND INTRODUCTION: This article reports evidence that the presentation order of audit working-paper documentation affects the outcome of audit partners' going-concern decisions. Although Frederick (1991) has found that memory organization and the organization of evidence interact to affect the retrieval of audit evidence from memory, this study establishes that memory organization and the organization of working-paper evidence interact to affect audit decisions.

Order effects are hypothesized because the going-concern decision task is driven by voluminous working-paper evidence that is typically documented in an order convenient for compiling audited financial statements, not in an order optimal for minimizing the cognitive effort required to reconstruct the causal order of relevant events. The experiment reported in this article tests the proposition that evidence documented in a contextually meaningful causal order (rather than in the traditional working-paper order of the public accounting firm participating in the study) will diminish the cognitive effort required of highly experienced audit partners by facilitating their ability to evoke pre-existing going-concern schemata from long term memory.

One hundred audit partners from a public accounting firm participated in the study. The experiment manipulated the presentation order of 60 sentences to compare the going-concern decisions of partners who read the audit evidence in variations of two distinctly different orders: causal order versus the participating firm's working-paper order. Evidence sentences were compiled from one of the firm's audit engagements by the researcher and by the engagement partner, manager, and reviewing partner assigned to the audit. Fifteen of the sentences referred to events consistent with a decision of substantial doubt about the entity's ability to continue as a going concern, 15 were consistent with a decision of no substantial doubt, 15 supported both decisions, and 15 supported neither

The author expresses appreciation to Larry Davis, Annemarie Keinath, Nancy Pennington, Greg Trompeter and the referees for insightful comments on earlier versions of this article.

Submitted June 1990.
Accepted August 1991.

decision. The sentences were arrayed in causal order by the researcher, engagement partner, and manager, using procedures outlined in Pennington and Hastie (1986, 1988). The working-paper order of presentation was prepared by the engagement partner, manager, and reviewing partner and represents the way an audit partner in the participating firm would likely encounter the evidence in a set of working papers assembled by the firm.

The experiment was constructed as a 2 x 2 design: 25 subjects read the "yes" sentences (evidence for substantial doubt) in causal order and the "no" sentences in working-paper order, 25 read the "no" sentences in causal order and the "yes" sentences in working-paper order, 25 read both the "yes" and "no" sentences in causal order, and 25 read both the "yes" and "no" sentences in working-paper order. To control for within-cell order effects, randomly assigned subjects in each of the four treatment groups read the "yes" cues first, and the others read the "no" cues first. All subjects completed two tasks: a decision regarding substantial doubt based on the 60 evidence sentences and an assessment of their confidence in the decision.

The results establish that the order manipulation affected the subjects' going-concern decisions but not their confidence in those decisions. Subjects decided there was substantial doubt about the client more often when the strongest evidence supporting this decision was read in causal order and, correspondingly, least often when the strongest evidence was read in the firm's working-paper order. The order effects documented in this study have important implications for audit working-paper design, particularly in complex unstructured decision tasks like the going-concern decision.

Key Words: *Audit decision making, Memory organization, Working papers, Order effects.*

Data Availability: *Data are available from the author upon request.*

THE remainder of this article is organized as follows. Section I discusses the hypothesized relationship between working-paper order and memory organization, and introduces the study's expectations. Section II documents the research method, and section III presents the results. Section IV summarizes the study and discusses implications for practice.

I. Working-Paper Order, Memory Organization, and Expectations

When making audit decisions, experienced auditors rely on a wealth of knowledge accumulated throughout their careers and on current audit evidence relevant to specific decision tasks. However, the process by which current evidence is encoded, stored, and retrieved from memory is directly dependent on how previously accumulated knowledge is organized in memory. For example, schema theory[1] proposes that what is stored in memory is a selected subset of the stimuli (e.g., audit evidence) presented to a decision maker, and his or her existing schemata are the "building blocks of cognition" (Rumelhart 1980, 33) that guide the selection process. As Ashton et al. (1988, 108) point out, the "role of schemata is particularly relevant in auditing, because the schemata developed by auditors through experience and prior knowledge of client situations may affect the manner in which the auditor perceives the evaluation of assertions and the need to accumulate and interpret evidence about the assertions."

---

[1] See Frederick (1991) for a discussion of a competing theory, taxonomic organization, and a comparison of a schematic and a taxonomic organization of knowledge in memory.

Underscoring the role of schemata in auditing, Frederick (1991) found that auditors freely recalled more internal controls when the controls were organized as a schematic representation (e.g., flow charts) rather than as a taxonomic representation (e.g., checklists). Frederick's results demonstrate that memory organization and the organization of audit evidence interact to affect the retrieval of evidence from memory. In contrast, the research reported here provides evidence that memory organization and the organization of working-paper evidence interact to affect auditors' decisions about the ability of an entity to continue as a going concern. Specifically, the study establishes that working-paper evidence presented in a contextually meaningful causal order has a greater influence on going-concern decisions than the same evidence presented in the usual working-paper order of the participating public accounting firm. The results provide initial evidence that the order in which evidence is documented in audit working papers could be an important source of differences in decisions across audit firms and, therefore, a potential source of decision error.

In a financial statement audit, an auditor assembles a large volume of evidence over several weeks or longer, and documents selected evidence in audit working papers. Most audit decision tasks are accomplished with evidence that is documented in adjacent working-paper pages, like tests of controls over sales transactions or substantive tests of accounts receivable. However, the going-concern decision, a task imbedded within the opinion-formulation process, is based on voluminous evidence documented throughout the audit working papers. Examples include evidence about results of operations, management turnover, debt-covenant performance, and contingent obligations, among other things. Owing to limitations of short-term memory, auditors rely on long-term memory to recall going-concern evidence encountered or read in working papers. Rereading working-paper documentation provides some relief from the demands on long-term memory (Libby 1989, 128), but the papers themselves often distort the temporal and causal order of relevant events. For example, the public accounting firm participating in this study normally groups evidence about loan covenants that restrict fixed-asset disposal toward the end of the "current working-paper file," and groups evidence about negative cash flows from operations, a potential causal reason for management's decision to dispose of fixed assets, at the beginning of a separate "working financial-statements file."

The participating firm assembles a current working-paper file in the order of a client's balance sheet and income statement account classifications, an order consistent with the way each account will appear in the audited financial statements. Because the firm's working papers are ordered for convenience in compiling financial statements, rather than in an order optimal for the going-concern decision, documented working-paper evidence is not likely to be presented in a manner that would diminish the cognitive effort required in a complex and unstructured going-concern decision.

To date, little is known about the role of working-paper order in audit decision making. For example, Cushing and Loebbecke (1986, chap. 4) include working-paper documentation among six variables used to proxy for the degree of audit structure in 12 public accounting firms, but the discussion relates primarily to the content, preparation, and administration of working-paper files, not to presentation order *per se*. Likewise, no one schema theory has emerged as the optimal predictor of memory organization in professional decision behavior. For example, Alba and Hasher (1983, 203-4) conclude that the literature has no fixed definition of the term "schema," reports largely conflicting experimental results, and proposes a sizeable number of paradigms (e.g., frame theory, script theory). Thus, no research in auditing or single theory in psychology provides a unique basis for predicting the effects of working-paper presentation order on auditors' going-concern decisions. As a result, I rely on characteristics of memory organization that are predicted by a variety of theories, among them schema theory (see Alba and Hasher 1983 for a review) and the more general notion of mental models, which account for the processing of atypical events that do not neatly fit preexisting schemata (Holland et al. 1986, 12–13).

In this article, I propose that voluminous audit evidence presented in a contextually meaningful causal order (as opposed to the participating firm's usual working-paper order) will diminish the cognitive effort required of highly experienced audit partners and facilitate their ability to evoke a preexisting going-concern schema from long-term memory. This proposition leads to two specific expectations. First, I expect that subjects receiving the strongest evidence for substantial doubt (about an entity's ability to continue as a going concern) in causal order, and the strongest evidence for no substantial doubt in working-paper order, are more likely to match the evidence with a consistent "yes" schema in memory and to conclude that there is substantial doubt. Because a causal schema is "one kind of pragmatic reasoning schema" (Holland et al. 1986, 78), this expectation is motivated by the assumption that causal order is a reasonable approximation of the memory organization and the preexisting going-concern schema of highly experienced audit partners. Subjects receiving evidence in the opposite order—the strongest "no" evidence in causal order and the strongest "yes" evidence in working-paper order—are less likely to conclude there is substantial doubt.

Second, I expect that the subjects' confidence in their decisions will be greater when evidence consistent with their decisions is presented in causal order rather than in working-paper order. That is, if causal order shifts decisions to the more easily constructed decision (the first expectation), then evidence presented in causal order is likely to appear more coherent to the subject, thereby increasing confidence in the decision. This expectation is motivated by the link between confidence and memory reported in auditing research. For example, Moeckel and Plumlee (1989) asked auditors to rate their confidence in relying on their own memories (rather than on a review of working papers and source documents) and found that auditors are at least as confident in their incomplete and inaccurate memories as they are in their accurate memories. Specifically, if confidence and memory are related, then significantly different confidence ratings across treatment groups could provide an insight into why the order manipulation shifted decisions. The experiment reported next tests both expectations.

II. Method

The design, expectations, and results of the experiment are summarized in table 1.

Table 1
Summary of Experiment

Dependent Variable: Substantial-doubt decision (yes vs. no)
Design: 2 (Order of "yes" evidence: causal/working paper) × 2 (Order of "no" evidence: causal/working paper)

Expectations	*Results*	*Results Support Expectation?*
Task: Substantial-doubt decision:		
Subjects who receive "yes" ("no") evidence in causal order and "no" ("yes") evidence in working-paper order are more likely to decide there is (not) substantial doubt.	Subjects' decisions in direction predicted and significant chi-square goodness-of-fit ($p<0.05$).	Yes
Task: Confidence rating:		
Subjects deciding there is (not) doubt are likely to have higher confidence ratings when "yes" ("no") evidence is presented in causal order and "no" ("yes") evidence is presented in working-paper order.	Subjects' confidence ratings in direction predicted but insignificant chi-square goodness-of-fit ($p>0.50$).	No

Subjects

Subjects for the experiment were 100 audit partners in a large public accounting firm. The participating partners, all of whom were selected in advance by the firm, had 11 to 32 years of public accounting experience (mean=18 years). Although the experiment provided no extrinsic incentives to the subjects, an intrinsic incentive was communicated by one of the firm's technical partners: the results of the experiment would be used in reconsidering the firm's prescribed working-paper design for financial statement audits. Despite the relatively broad range of experience among the subjects, differences in going-concern decisions were not predicted (or observed) for subsets of the partners, since all of them maintain at least some audit client responsibilities and, therefore, are required under SAS No. 59 (AICPA 1988) to make (or approve) going-concern decisions for each of their audit clients.

Materials

The experimental session was held in a microcomputer lab at an annual partners' meeting. Each subject was seated at a display monitor and told that he or she would view 60 evidence sentences, one sentence at a time, and then make a going-concern decision. The subjects worked individually and at their own pace, advancing to the next sentence by depressing the keyboard spacebar (following display, none could be retrieved by the subjects).[2] Although the subjects received neither financial statements nor background narrative, the evidence sentences included both financial cues (e.g., financial results) and background (e.g., products, top management changes) relevant to the decision.[3]

---

[2] Response times were not compiled since prior research suggests no *a priori* reason to expect differences across subjects (Pennington and Hastie 1988).

The 60 evidence sentences were derived from a set of audit working papers by the researcher and by three representatives of the participating firm: the engagement partner, manager, and reviewing partner assigned to the audit. Although the client was heavily burdened by debt obligations and poor cash flow, the firm's representatives had judged the case "at the margin," one that could steer the subjects toward either decision, yes or no.

The three representatives of the participating firm reviewed the working papers independently, identifying and ordering causally connected evidence descriptors that supported or diminished substantial doubt about the client's ability to continue as a going concern. The representatives agreed on 85 percent of the descriptors and resolved differences by group discussion. This evidence and other relevant, though not causally connected, information was transformed into brief statements that were classified by the researcher and the engagement partner according to whether the statements supported (1) the "yes" decision for substantial doubt about continued existence, (2) the "no" decision, (3) both decisions, or (4) neither decision. The intent was to avoid constructing any one strongly persuasive sentence that might alone drive premature decisions, for example, "The company plans to file for bankruptcy." Premature decisions could have detracted the subjects' attention from the remaining sentences. (Section III reports a pilot study that was conducted, in part, to facilitate a test of recency and primacy in the experiment.)

Fifteen of the 60 sentences referred to events consistent with the "yes" decision, for example, "The company will report a net loss for the second straight year." Fifteen referred to events consistent with the "no" decision, for example, "The company entered into a long term contract to supply copper tubing to a new domestic customer." Fifteen referred to events consistent with both decisions, for example, "The company contemplates issuing bonds next year."[4] And 15 referred to events consistent with neither decision, for example, "Independence is not impaired by any staff assignments."

An appendix lists all 60 sentences and their classification according to the decision each sentence supports (i.e., yes, no, both, neither). The appendix also indicates both working-paper order and causal order. Working-paper order was constructed by the engagement partner, manager, and reviewing partner, and reflects the way an audit partner would likely encounter the evidence in a set of working papers assembled by the firm. Causal order was constructed by the researcher, engagement partner, and manager according to procedures outlined in Pennington and Hastie (1986, 1988).

Although the pattern of results predicted in this article are also predicted by other schema theories and by mental models, Pennington and Hastie's model was selected to construct the causal order of the evidence sentences for two reasons. First, the model is sufficiently robust to capture audit settings because the going-concern decision is a prototype of the tasks for which the model predicts order effects: An auditor makes a cost-constrained decision from a large data base of interdependent evidence that is piecemeal in its depiction of historical events and that is received in unordered sequence over a period of time. Second, the model has successfully predicted decision behavior in experiments that have used both traditional quasi-experimental methodologies (Pennington and Hastie 1988) and process tracing methodologies (Pennington and Hastie 1986).

---

[3] Since the experimental materials were adapted from an actual audit client, the engagement partner and the reviewing partner felt that actual background information may have revealed the client's identity and that hypothetical information may have distorted the task's realism.

[4] The sentence refers to both the "yes" and "no" decisions in this sense: to the "yes" decision because offering bonds results in additional debt (and, therefore, additional demands on cash for interest and principal payments), to the "no" decision because the bond proceeds provide a means to finance existing liabilities.

A causal structure was constructed separately for the 30 statements consistent with substantial doubt and the 30 sentences consistent with no substantial doubt. In the causal structure, each of the evidence sentences is classified according to two criteria (Pennington and Hastie 1988, 525): the sentence's degree of causal connectedness (the number of direct causal connections with all other sentences) and whether the sentence is on the causal chain (the subset of statements that the engagement partner and manager judged most persuasive to the going-concern decision in this case). The causal chain for the "yes" decision consists of the causal-order sentence numbers 7, 12, 13, 14, 16, 19, 20, 29, and 50 in the appendix. The causal chain for the "no" decision consists of sentence numbers 17, 18, 28, 33, 34, 37, 42, 45, and 58.

Design

The experiment was constructed as a 2×2 design: 25 subjects read the "yes" sentences in causal order and the "no" sentences in working-paper order (cell 1), 25 read the "no" sentences in causal order and the "yes" sentences in working-paper order (cell 2), 25 read both the "yes" and "no" sentences in causal order (cell 3), and 25 read both the "yes" and "no" sentences in working-paper order (cell 4). To remove within-cell order effects, randomly assigned subjects in each of the four cells read the "yes" cues first, and the remaining subjects read the "no" cues first. For example, in cell 1, 13 randomly selected subjects (i.e., about half) read the 30 "yes" statements (15 "yes" plus 15 "both") in causal order first and then read the other sentences (15 "no" plus 15 "neither") in working-paper order; in contrast, the remaining 12 subjects read the working-paper order sentences first and then read the causal order sentences. In cell 2, 13 randomly selected subjects read the 30 "no" statements (15 "no" plus 15 "both") in causal order first and then the other sentences (15 "yes" plus 15 "neither") in working-paper order; the remaining 12 subjects read the working-paper order sentences first and then read the causal order sentences.[5]

Tasks

After reading all 60 evidence sentences in succession at a display monitor, each of the 100 subjects completed the first task: indicate yes there is or no there is not substantial doubt about the entity's ability to continue as a going concern within one year of the balance sheet date, the task imposed by SAS No. 59 (AICPA 1988). If the first expectation (table 1) in the study is correct, the number of subjects deciding that there is substantial doubt should be highest for the subjects who read the "yes" sentences in causal order and the "no" sentences in working-paper order (cell 1), and lowest for the

---

[5] There was no significant difference in within-cell mean confidence ratings attributable to the item order manipulation ($p > 0.10$).

Table 2
Substantial-Doubt Decisions and Mean Confidence Ratings Classified by Evidence Presentation Order

	Presentation Order		*Substantial-Doubt Decision and Mean Confidence Ratings*				
			Yes		*No*		
Cell	*Yes Sentences*	*No Sentences*	*n*	*Rating*	*n*	*Rating*	*Mean Cell Rating**
1	Causal	WP	21	3.82	4	3.15	3.49
2	WP	Causal	8	3.13	17	3.71	3.42
3	Causal	Causal	14	3.46	11	3.20	3.33
4	WP	WP	15	3.22	10	3.18	3.20
			58	3.41	42	3.31	3.36

*Unweighted arithmetic mean
WP = Working-paper order

subjects who read the "no" sentences in causal order and the "yes" sentences in working-paper order (cell 2).

In the second task, subjects assessed their confidence in their decision on a five-point scale (1 = pure guess, 5 = very certain). If the second expectation (table 1) in the study is correct, then subjects deciding there is doubt are expected to have higher confidence in their decisions when the "yes" sentences are presented in causal order and the "no" sentences in working-paper order, rather than when the "no" sentences are in causal order and the "yes" sentences in working-paper order. Likewise, subjects deciding against substantial doubt are expected to have higher confidence when the sentences are presented in the opposite order: "no" sentences in causal order and "yes" sentences in working-paper order.

III Results

Table 2 reports the subjects' substantial-doubt decisions and mean confidence ratings classified by the presentation order of the evidence sentences. The pattern of "yes" (and therefore "no") decisions across cells in table 2 confirms the study's first order-effect expectation. That is, subjects were more likely to decide there is doubt when the strongest evidence supporting this decision ("yes" sentences) was read in causal order, and the strongest evidence against ("no" sentences) was read in working-paper order (cell 1: n = 21). Correspondingly, subjects were least likely to decide "yes" when "no" sentences were read in causal order and "yes" sentences in working-paper order (cell 2: n = 8). The subjects who read both the "yes" and "no" sentences in causal order or in working-paper order were almost equally likely to decide there is doubt (cell 3: n = 14; cell 4: *n* = 15). The pattern of decisions in cell 1 (21 yes, 4 no) and cell 2 (8 yes, 17 no) suggests that the order manipulation in the experiment shifted the subjects' decisions in the direction of the more easily constructed decision. If the subjects' decisions were independent of the order manipulation, the expected number of "yes" decisions within cells 1 and 2 would be roughly proportional to the percentage of "yes" decisions in the experiment: 14.5 per cell (58 of 100 subjects decided "yes": $58/100=0.58\times 25$ subjects per cell$=14.5$), as it was in cells 3 (14 yes) and 4 (15 yes). A chi-square goodness-of-fit test leads to rejection of the null hypothesis that the reported "yes" decisions do not differ significantly from 14.5, the expected number ($X^2(1, n=100)=6.80$, $p<0.5$), which confirms the expectation that the subjects' decisions were not likely independent of the experiment's order manipulation. Note, however, that because the order of the "both" and "neither" sentences changed when the causal and working-paper orders of the "yes" and "no" sentences were manipulated across treatment groups, the design does not address whether the reported order effects are driven by processing differences related to the yes-no stimuli or the both-neither stimuli.

The pattern of mean confidence ratings within cells 1 and 2, the two cells most relevant to the study's confidence expectations, is in the expected direction. That is, subjects deciding "yes" there is doubt had higher confidence ratings when "yes" sentences were presented in causal order and "no" sentences in working-paper order (cell 1: 3.82) than when "no" sentences were in causal order and "yes" sentences in working-paper order (cell 2: 3.13). Likewise, subjects deciding "no" had higher confidence when "no" sentences were presented in causal order and "yes" sentences in working-paper order (cell 2: 3.71) than when "yes" sentences were in causal order and "no" sentences in working-paper order (cell 1: 3.15). However, a chi-square goodness-of-fit test failed to reject the null hypothesis that the subjects' confidence ratings across cells do not differ significantly from the grand mean (3.36) at table 2 ($X^2(3, n=100)=0.15$, $p>0.50$). Thus, the results do not confirm the expectations about confidence.

In this experiment, the order-effect results are subject to the alternative explanation that the subjects' decisions may have been driven by the recency or primacy of selected sentences. Recency and primacy effects were tested by using a two-step approach. First, in a pilot study, 40 additional audit partners read the 60 evidence sentences in different random orders and then performed two tasks: a substantial-doubt decision (yes/no),[6] and a rating of the importance each sentence had for their decision on a 65-point scale. The scale was anchored by "Important evidence *against* substantial doubt" and "Important evidence *for* substantial doubt," the midpoint being "Not important." To test for recency and primacy effects, correlations were calculated between the presentation order of each sentence in the experiment and its mean importance rating from the pilot study. None of the correlations was significantly different from zero ($p>0.10$), which suggests that the subjects' decisions in the experiment were not likely biased by the location of evidence presented.

---

[6] Apart from facilitating a test of recency and primacy, the pilot study was also used to gather evidence about whether the experimental materials would likely drive different going-concern decisions among the partners participating in the actual experiment. The 40 audit partners participating in the pilot study ranged in public accounting experience from 13 to 30 years (mean = 17 years), an experience level roughly equivalent to the 100 partners participating in the experiment. In the pilot study, 24 of the 40 partners decided "yes" there is substantial doubt and 16 decided "no." Because no two pilot subjects read the sentences in exactly the same order, the results of the pilot study could not be attributed to order effects, suggesting that the evidence sentences in the experimental materials would likely be successful in creating a substantial-doubt decision task that was "at the margin."

IV. Summary and Discussion

The experiment reported in this article provides evidence that the presentation order of evidence sentences affected audit partners' substantial-doubt decisions (although not their confidence in those decisions). The subjects decided "yes" there is substantial doubt more often when the strongest evidence supporting the decision was presented in causal order and, correspondingly, less often when the strongest supporting evidence was presented in the order customarily found in the participating firm's audit working papers. Although recent research reports differences in knowledge structures between groups of experienced and inexperienced auditors in recall and inference tasks (Choo and Trotman 1991), no prior work provides evidence of an interaction between the organization of knowledge structures in memory and the organization of documented audit evidence within groups of highly experienced audit partners. Future research should focus on the external validity of the order effects documented here, the objective being to stage optimal presentation orders for specific audit decision tasks.

Although not generalizable beyond the participating firm, the results of this study are potentially significant to audit practice because they have implications for the order in which audit evidence is arranged for decision making. Little information is available publicly about the design of public accounting firms' proprietary audit working papers. However, to the extent that firms order their working papers similar to the participating firm (i.e., in the order of balance sheet and income statement accounts), then this research suggests that there is reason for firms to consider the potential for order effects in unstructured decision tasks that are based on voluminous evidence, particularly when balance sheet and income statement order differs significantly from causal order for the task.

Despite the ready availability of working-paper software (see Hoffman 1988 for a review), many public accounting firms (including the participating firm) continue to store audit evidence as hard copy documents partitioned by financial statement accounts. The results of this study suggest it may be advantageous for firms to store audit evidence in automated data files that function much like database management systems. For example, in the going-concern decision, evidence could be gathered by traditional audit methods, stored in automated databases, and accessed by expert systems software that displays the evidence in a sequence that preserves the causal order of events. Database working papers could minimize the interactive effect of cue order on memory organization, allowing the auditor to focus attention on cue selection and cue weighting, two of the more telling issues in complex decision making.

Appendix
Evidence Sentences
(Decision Classifications in Parentheses)

*Order**		
C	WP	*Evidence Sentence*
Background		
1	13	The company was incorporated in 1955. (Neither)
2	14	The company's common stock is traded over the counter. (Neither)
3	15	The company manufactures wire and cable products. (Neither)
4	1	The engagement letter called for an audit opinion and preparation of Form 10K. (Neither)
5	2	Independence is not impaired by any staff assignments. (Neither)
6	3	The company imposed no restrictions on the nature, timing, or extent of audit procedures. (Neither)
Changes in Management		
7	16	Top management has changed in the last year. (Yes)
8	4	The president announced he was considering a competitor's employment offer. (Yes)
9	5	The resignation of the vice-president for manufacturing is effective in six months. (Yes)
Labor Relations		
10	7	In midyear, there was a two-day work stoppage related to employee fringe benefits. (Yes)
11	57	A newspaper clipping indicates the company's production employees voted not to strike. (Both)
Demand/Market Share		
12	18	Domestic demand for wire and cable products is down. (Yes)
13	55	Competition from foreign imports has risen sharply. (Yes)
14	42	A patent for a unique production process expires next year. (Yes)
15	58	A principal customer has not placed an expected order. (Yes)
16	56	A trade journal noted the company's market share had declined. (Yes)
17	6	The company agreed in substance to ship copper products to a Pacific rim country. (No)
18	34	The company entered into a long-term contract to supply copper tubing to a new domestic customer. (No)
19	43	The company was denied credit from a copper supplier. (Yes)
20	35	The company entered into an uneconomic long-term purchase commitment for copper. (Yes)
Research, Development, and Capacity		
21	8	The company plans a 50 percent reduction in budgeted research and development expenditures. (Both)
22	30	The company postponed a consulting project related to new product development. (Both)
23	36	The company is operating at 70 percent of capacity. (Both)
Risk and Materiality		
24	17	Inherent risk is low. (Neither)
25	27	Control risk is low. (Neither)
26	28	Allowable detection risk is high. (Neither)
27	29	The preliminary estimate of materiality is moderate. (Neither)
Financial Results		
28	20	Revenues rose 5 percent over last year. (No)
29	23	The company will report a net loss for the second straight year. (Yes)
30	19	Two competitors reported net losses last year. (Neither)
31	21	There is no significant change in total assets since last year. (No)
32	41	The current ratio is lowest in the industry. (Yes)
33	31	The number of days sales in receivables is lower than in any of the previous five years. (No)
34	32	There are no trade-receivable accounts more than 90 days old. (No)
35	33	No trade receivables have been factored. (Both)
36	22	There is no significant change in total liabilities and equities since last year. (No)
37	50	The debt-to-equity ratio is down this year. (No)

Appendix—*Continued*

Order*		
C	WP	Evidence Sentence
Results of Tests of Controls and Substantive Tests		
38	26	Tolerable rates of deviation exceeded maximum population rates for all tests of controls. (Neither)
39	53	Substantive tests were performed for all financial statement assertions. (Neither)
40	54	Tolerable error exceeded projected monetary error for all financial statement accounts. (Neither)
Tangible Assets		
41	24	The company plans to delay scheduled maintenance on unused fixed assets. (Both)
42	37	Unused fixed assets are readily marketable. (No)
43	52	Loan covenants restrict the disposal of some fixed assets. (Yes)
44	38	The replacement cost of fixed assets is not fully insured. (Both)
Debt and Equity		
45	51	The company has no bond indebtedness outstanding. (No)
46	59	The company contemplates issuing bonds next year. (Both)
47	60	The company contemplates an initial public offering of preferred stock next year. (Both)
48	11	A dividend has not been declared in four consecutive quarters. (Both)
49	12	The company sought to finance the purchase of a subsidiary that had significant NOL carry-forwards. (Both)
Bank Loans		
50	44	The company defaulted a payment on its only bank loan. (Yes)
51	45	The company renegotiated its only bank loan. (Both)
52	9	A loan from a director has been forgiven. (Both)
53	39	There are no arrangements with financial institutions for compensating balances. (No)
54	40	A confirmation indicates the company is not contingently liable to the bank. (No)
Insurance		
55	10	A personal injury insurance policy was cancelled. (Both)
56	25	The company has never filed an insurance claim. (Neither)
Litigation, Claims, and Assessments		
57	46	There are no violations of laws requiring disclosure as a loss contingency. (No)
58	47	The company has complied with all contractual agreements. (No)
59	48	A lawyer's letter indicates no current litigation. (No)
60	49	A newspaper clipping indicates the company may be sued for patent infringement. (Both)

* C = Causal order.
WP = Working-paper order.

References

Alba, J. W., and L. Hasher. 1983. Is memory schematic? *Psychological Bulletin* 93 (March): 203–31.

American Institute of Certified Public Accountants. 1988. *Statement on Auditing Standards No. 59*, The auditor's consideration of an entity's ability to continue as a going concern. New York: AICPA.

Ashton, A. H., D. N. Klienmuntz, J. B. Sullivan, and L. A. Tomassini. 1988. Audit decision making. In *Research Opportunities in Auditing: The Second Decade*, edited by A. Rashad Abdel-khalik and I. Soloman, Sarasota, FL: AAA, 95–132.

Choo, F., and K. T. Trotman. 1991. The relationship between knowledge structure and judgments for experienced and inexperienced auditors. *The Accounting Review* 66 (July): 464–85.

Cushing, B. E., and J. K. Loebbecke. 1986. *Comparison of Audit Methodologies of Large Accounting Firms.* Sarasota, FL: AAA

Frederick, D. M. 1991. Auditors' representation and retrieval of internal control knowledge. *The Accounting Review* 66 (April): 240–58.

Hoffman, R. B. 1988. AICPA computer survey. *Journal of Accountancy* 166 (September): 146–50.

Holland, J. H., K. J. Holyoak, R. E. Nisbett, and P. R. Thagard. 1986. *Induction: Processes of Inference, Learning, and Discovery.* Cambridge, MA: MIT Press.

Libby, R. 1989. Experimental research and the distinctive features of accounting settings. In *Illinois PhD Jubilee 1939–1989: The State of Accounting Research As We Enter the 1990s,* edited by T. J. Frecka, Urbana-Champaign: University of Illinois, 126–47.

Moeckel, C. L., and R. D. Plumlee. 1989. Auditors' confidence in recognition of audit evidence. *The Accounting Review* 64 (October): 653–66.

Pennington, N., and R. Hastie. 1986. Evidence evaluation in complex decision making. *Journal of Personality and Social Psychology* 51 (August): 242–58.

———, and ———. 1988. Explanation-based decision making: The effects of memory structure on judgment. *Journal of Experimental Psychology: Learning, Memory and Cognition* 14 (July): 521–33.

Rumelhart, D. E. 1980. Schemata: The building blocks of cognition. In *Theoretical Issues in Reading Comprehension,* edited by R. J. Spiro et al. Hillsdale, NJ: Lawrence Erlbaum.

THE ACCOUNTING REVIEW
Vol. 67, No. 1
January 1992
pp. 59–76

Explanation and Counterexplanation During Audit Analytical Review

Lisa Koonce
The University of Texas at Austin

SYNOPSIS AND INTRODUCTION: Analytical review, the diagnostic process of identifying and determining the cause of unexpected fluctuations in account balances and financial ratios, has become an increasingly important part of auditing. Recent research suggests that analytical review is beneficial for detecting such unexpected fluctuations (Kreutzfeldt and Wallace 1986; Wright and Ashton 1989). It is, however, difficult for an auditor utilizing analytical review to know when the "correct" cause for an observed fluctuation has been determined, especially since no official guidance exists to assist the auditor. Thus, accepting a plausible but incorrect cause is a dangerous and largely unresearched possibility (Kinney 1987).

This article describes two experiments that investigate conditions under which auditors may be prone to accepting a plausible, but potentially incorrect, cause when performing attention-directing analytical review. Identifying such conditions is important because either audit effectiveness or efficiency would be compromised with such incorrect acceptance. Although there are no normatively correct causes for the analytical review problem situations employed in this study, relative evidence on such compromises will be obtained by investigating conditions that lead auditors to make different judgments about the cause of an unexpected fluctuation. Differences in auditors' judgments are important since they can lead to differences in decisions as to the cause of an unexpected fluctuation and subsequent auditor actions (cf. Asare 1991).

Dichotomizing causes into error and non-error classes reveals several

---

I give special thanks to Ira Solomon, my dissertation chairman, for his assistance with this project. I also thank Urton Anderson, Clif Brown, Bill Kinney, Garry Marchant, Jim McKeown, two anonymous referees, and an anonymous associate editor for their helpful comments on a previous version of this article. Also, the workshop participants at Duke University, Ohio State University, University of Illinois, University of Texas, and University of Arizona provided many useful comments. Funding was provided by the University of Illinois and by the Ernst & Whinney and Touche Ross & Co. Foundations.

Submitted July 1990.
Accepted June 1991.

ways in which an auditor performing analytical review may accept a plausible but incorrect cause. One that has particularly significant implications for the effectiveness of the audit would be concluding that a *non-error* cause adequately accounts for an unexpected fluctuation when the "correct" cause involves a financial-statement *error.* Since it is unlikely that the auditor would increase planned tests of details if indications of a material error were not present (and it is possible that the auditor may decrease such testing), the likelihood of subsequent detection of the error would be reduced in this situation. Conditions that may lead to this type of decision error will be investigated in this study.

The effects of explanation, counterexplanation (i.e., consideration of why an hypothesized cause may not be correct), and their order of performance were studied to determine how such factors affect the propensity for acceptance of a plausible but incorrect non-error cause. Investigating explanation is important since auditors typically document explanations for analytical review findings. Even though current auditing standards do not require auditors to counterexplain, investigating how counterexplanation and the order of counterexplanation and explanation affect auditors' judgments is nevertheless important since it enhances our understanding of the potential for acceptance of an incorrect cause.

Auditors from five public accounting firms participated in two experiments conducted to determine the effects of the aforementioned factors on auditors' judgments. Consistent with expectations, auditors revised their beliefs about an hypothesized non-error cause of an unexpected fluctuation in the auditee's gross margin in an upward (downward) fashion when asked to document an explanation (counterexplanation) for that cause. Inconsistent with expectations, however, auditors who were asked to document an explanation initially and then a counterexplanation judged the non-error cause to be less likely to have occurred than did auditors asked to document a counterexplanation initially and then an explanation. Overall, these results identify conditions in which auditors may be prone to compromising audit effectiveness by accepting a potentially incorrect non-error cause.

Key Words: *Analytical review, Explanation, Counterexplanation.*

Data Availability: *Data used in this study are available from the author upon request.*

THIS article is organized as follows. Section I describes how three factors—explanation, counterexplanation, and the order in which these occur—can affect the potential for acceptance of a plausible but incorrect cause during analytical review. The theory and hypotheses also are developed in this section. Section II provides a description of the method and results of the first experiment, which investigates the impact of explanation on auditors' analytical review judgments. Section III presents the method and results of the second experiment, which investigates the impact of counterexplanation and task-performance order (i.e., explanation preceding counterexplanation vs. counterexplanation before explanation) on auditors' analytical review judgments. The final section provides a discussion of the results and concluding comments.

I. Theory and Hypotheses

Analytical review is characterized by sequential and diagnostic causal reasoning that frequently involves the development of written explanations (Gibbins 1984; Libby 1985). An appropriate theoretical framework for studying incorrect acceptance of a cause within analytical review should recognize these contextual characteristics. Belief perseverance, a social-cognition theory, predicts that when information is *sequentially* and *causally* evaluated, information considered first will have a greater effect on an individual's final (revised) belief than information considered later. Because of its ability to accommodate the aforementioned analytical review characteristics, belief-perseverance theory was chosen to guide the investigation of the potential for acceptance of a plausible but incorrect non-error cause.[1]

The Explanation Effect

When performing analytical review at the planning stage of the audit, the auditor initially determines an expectation for the account balance or financial ratio under investigation. This expectation assists the auditor in identifying any unexpected fluctuation. Once such a fluctuation has been identified, the auditor often will secure a potential cause from client management (AICPA 1988a, par. 21; Wright and Ashton 1989) and, at some point, may document an explanation supporting that cause in the audit workpapers. Because information supporting the validity of the inherited cause must be considered when explanations are required, such explanations should increase the extent of causal reasoning employed by the auditor relative to situations in which written explanations are not required.

Prior belief-perseverance research has found that the increased causal reasoning accompanying explanation development typically causes an event that is part of a written explanation to be judged more likely to have occurred than an event that is not part of a written explanation. This phenomenon has been termed an explanation effect (Anderson and Sechler 1986). Observing this phenomenon is important because individuals who develop written explanations are more likely to maintain their existing beliefs even when subsequent data challenge those beliefs than are individuals who have not developed written explanations (Anderson et al. 1980).

The foregoing cognitive mechanism is congruent with the auditor's sequential belief-revision process for analytical review. That is, since auditors are required by professional standards to document explanations for audit findings (AICPA 1982), an explanation effect is expected when they develop written explanations for an hypothesized cause of an unexpected fluctuation. If an explanation effect is observed, it would suggest that auditors who write explanations would be more prone to accepting an incorrect cause than auditors who do not write such explanations.

This prediction is made in light of prior research which has reported that the explanation effect does *not* occur with subjects experienced in the task. Specifically, Anderson and Wright (1988) report that for an internal-control evaluation task, auditors' judgments about the likelihood of various errors and irregularities are not affected by whether auditors explain how those events could happen. Although these results suggested to Anderson and Wright that auditors have stored in memory previously generated causal scenarios for these events, one aspect of the experimental procedure may provide another interpretation. In particular, as part of developing their written explanations, subjects were asked to provide reasons both supporting and opposing the occurrence of the target event. Psychology research has shown that the generation of opposing reasons (i.e., counterexplanation) will negate any observed explanation effect, which is typically measured by asking subjects only to provide supporting reasons (Anderson and Sechler 1986). Hence, it remains an open question as to whether Anderson and Wright would have obtained an explanation effect (as typically measured in the psychology literature) had they not asked their subjects to counterexplain.

---

[1] As used herein, non-error causes refer to unusual transactions or events, accounting changes, or random fluctuations; error causes refer to intentional and unintentional financial-statement misstatements.

To determine whether the explanation effect will occur in the audit domain necessitates an experimental task in which subjects are requested to engage in explanation only and the potential for counterexplanation is minimized. The task of experiment 1 was designed to meet these criteria. Consequently, based on psychology research suggesting that an explanation effect will occur when auditors document explanations for unexpected fluctuations during analytical review, the following alternative-form hypothesis is proposed:

H1: Auditors who are asked to generate a written explanation for an hypothesized non-error cause will assign a higher probability to that cause than auditors who are not asked to generate a written explanation.

The Counterexplanation Effect

The causal reasoning that occurs when written explanations are developed should enable the auditor to acquire a more reasoned judgment about the validity of the hypothesized cause. However, it also may cause the auditor to prematurely cease information search and accept a seemingly plausible but incorrect cause. Prior belief-perseverance research has shown that individuals seldom spontaneously counterexplain (i.e., consider information not supporting a target event) once a written explanation has been developed (see, e.g., Anderson 1982; Anderson and Sechler 1986). The extent to which individuals engage in counterexplanation, however, is greater in situations in which they are requested to develop a written counterexplanation. Information *not* favoring a target event becomes more developed and salient in memory for those who document such counterexplanations. The latter enhances the perceived multiplicity of causation, thus attenuating belief in the validity of the target event (Einhorn and Hogarth 1986). This reduction in the judged likelihood of a target event once a written counterexplanation is developed is called a counterexplanation effect and is important because individuals who document counterexplanations are less likely to accept the implications of an explained target event than are individuals who do not document counterexplanations.[2]

---

[2] Although it is possible that the change in the assessed probability of the non-error cause created by a written explanation (or counterexplanation) may not be significant if an auditor's actions are invariant over the affected range, establishing that there is a change in probability which suggests that at least some auditors will cross a decision/action threshold. This is especially likely if one assumes that there are many different thresholds resulting from many different potential action choices.

Extant research has investigated the counterexplanation effect in situations conducive to the phenomenon, that is, in situations in which individuals were not expected to counterexplain spontaneously (e.g., asking college students to consider whether firefighters are risk takers [Anderson 1982]). The propensity for spontaneous counterexplanation may be heightened in the audit environment in light of professional standards that require the auditor to maintain an attitude of professional skepticism during the conduct of the audit (AICPA 1988b). Although professional skepticism is not procedurally defined in the standards, exhibiting such skepticism could arguably involve the process of counterexplaining (Pei et al. 1991).

Heiman (1990) has investigated the issue of whether auditors spontaneously counterexplain when performing audit analytical review, but her experimental task appears to have involved a scenario in which spontaneous counterexplanation was not likely. Specifically, Heiman's subjects were given an error cause by the *researcher* in response to an observed fluctuation. Since her subjects may have attributed some reason to the provision of that error cause (e.g., the researcher knew that the error cause was correct), their predisposition to counterexplain spontaneously may have been diminished. In this study, subjects' professional skepticism will be heightened by providing them with a *non-error* cause from *client management*. Client management's incentives to offer non-error, rather than error, reasons in response to auditor inquiries should increase the propensity for spontaneous counterexplanation. Further, Heiman's subjects were given no counterexplanation information in the case materials, thus requiring them to generate such information from memory. These factors, in conjunction with the subjects being told that there was only *one* error cause for the observed fluctuation, may have further reduced any predisposition to counterexplain spontaneously, thus leading to the observed counterexplanation effect. Subjects in this study will be *provided* counterexplanation information in the case materials and told that *multiple* causes are possible, thereby enhancing the likelihood of spontaneous counterexplanation.[3]

Belief-perseverance theory does not indicate the extent to which factors promoting professional skepticism in the audit context will increase the propensity for spontaneous counterexplanation. The purpose of this study, therefore, is to determine whether a counterexplanation effect will occur in a context in which there is an increased likelihood of spontaneous counterexplanation. Accordingly, this study will assist in identifying boundary conditions on the counterexplanation effect by providing a stronger test than yielded from prior research. However, based on our present state of knowledge, which suggests that individuals do not spontaneously counterexplain (or at least not as much as when explicitly instructed to do so), the following alternative-form hypothesis is proposed:

H2: Auditors who are asked to develop a written explanation for an hypothesized non-error cause will assign higher probabilities to that cause than auditors who are asked to develop a written explanation *and* a written counterexplanation.

---

[3] Another difference between this study and Heiman's involves the theory upon which the experimental predictions are based. Specifically, Heiman used a discounting model rather than the theory, belief perseverance, used herein. This theory has important advantages over the discounting model, including its ability to predict how the development of a written explanation affects the propensity for spontaneous counterexplanation. Lacking this, Heiman was unable to investigate the effect of explanation on spontaneous counterexplanation.

The Primacy Order Effect

Although development of a written counterexplanation is expected to reduce the propensity to accept the implications of an explained target event, various approaches to counterexplanation have been suggested to differentially reduce this propensity. In particular, research indicates that the order in which explanation and counterexplanation procedures are employed can differentially affect target-event probability assessments (Hoch 1984; Moser 1989; cf. Tetlock 1983). When task-performance order is explanation followed by counterexplanation, belief-perseverance theory predicts that final probability assessments regarding the occurrence of the target event will be higher than if task-performance order were counterexplanation then explanation. This pattern of results is consistent with a primacy order effect (cf. Anderson 1981) and suggests that auditors who explain and then counterexplain will be more prone to accepting an incorrect cause than will auditors who follow the opposite task-performance order.

The primacy order effect occurs because initially developing a target-event explanation frequently causes an individual to overweight that explanation relative to the opposite task-performance order (Nisbett and Ross 1980; Tetlock 1983). In some cases, such initial generation of an explanation can also cause the individual to be less able or willing to subsequently generate counterexplanation arguments (Anderson et al. 1985; Hoch 1984; Moser 1989). In contrast, when information not favoring a target event is initially considered, uncertainty surrounding the validity of the target event is greater since the perceived multiplicity of causation has been enhanced *before* a written explanation is constructed (Anderson 1982). Consequently, individuals who initially counterexplain and then explain tend to overweight (underweight) their counterexplanations (explanations) relative to individuals employing the opposite task-performance order.

In conformance with belief-perseverance theory, auditors who initially engage in explanation followed by counterexplanation are expected to assess higher final probabilities for a non-error cause (and, thus, be more prone to accepting an incorrect cause) than auditors who initially engage in counterexplanation followed by explanation. This primacy order effect is predicted even though other auditing research has consistently reported *recency* order effects for auditors' sequential processing of information (see, e.g., Ashton and Ashton 1988; Tubbs et al. 1990). Recency occurs when information has a greater impact on beliefs when it is evaluated last rather than first in a two-task sequence.

The theoretical framework used for these audit studies was Hogarth and Einhorn's (forthcoming) belief-adjustment model. This model predicts recency, primacy, or no-order effects, depending on various contextual factors. However, a basic premise of the Hogarth and Einhorn model makes it incongruent with audit analytical review and, thus, precludes its use in this study. Specifically, the model is appropriate only for experimental contexts in which all information items are presented to subjects who are not allowed to generate explanations or otherwise integrate the information items. Since belief-perseverance theory recognizes the importance of integrating information items and developing mental and written explanations in the judgment-formulation process, it provides a more realistic portrayal of analytical review which involves such cognitive processes. Because belief-perseverance appears to be more suitable than the belief-adjustment model for studying order effects within audit analytical review, the following alternative-form hypothesis is proposed:[4]

> H3: Auditors who initially are asked to develop a written explanation and then develop a written counterexplanation will assess higher probabilities for the hypothesized non-error cause than auditors who initially are asked to develop a written counterexplanation and then develop a written explanation.

Two experiments test the three hypotheses. The purpose of experiment 1 is to determine how developing a written explanation for a non-error cause affects auditors' probability assessments for that cause (hypothesis H1). The purpose of experiment 2 is to determine how developing a written counterexplanation affects the probabilities assigned to the non-error cause (hypothesis H2) and how task-performance order also affects these probability judgments (hypothesis H3). Two experiments are necessary since the case materials for testing hypotheses H2 and H3 must include counterexplanation information while the case materials for testing hypothesis H1 must exclude such information. Such exclusion is necessary to avoid the possibility of the subjects spontaneously incorporating counterexplanation evidence into their judgments and, thus, masking the effects of explanation.

II. Experiment 1—The Explanation Effect

Materials and Subjects

There were three components to the case materials. Part A described a hypothetical first-year manufacturing client to which subjects were assigned the task of performing attention-directing analytical review. In addition, part A contained financial and non-financial background information about the client and its industry (the computer industry).[5] Part B of the case materials contained narratives summarizing prior discussions with various client personnel. These narratives contained information pertaining to (but not perfectly diagnostic of) the non-error cause (a change in sales mix) that the client subsequently would suggest led to an unexpected fluctuation in the gross margin. Two different text orders of the discussion narratives were utilized to avoid the potential influence of a particular text order on the results. Finally, part C of the case materials contained either four or five questions (including a post-experimental questionnaire), depending on the experimental condition.

Participants were 34 senior-level auditors from five public accounting firms with experience ranging from 1.7 to 5.0 years (average experience of 2.8 years). The use of audit seniors is justified, since they typically perform the attention-directing analytical review task employed in this study and have reasonably well-developed knowledge of potential (error and non-error) causes of unexpected fluctuations (Libby and Frederick 1990). The experiment was administered to the subjects (on a group basis by firm) in their offices during regular business hours. Subjects were not allowed to use reference

[4] Even if the belief-adjustment model were deemed more appropriate for describing how auditors evaluate causes during analytical review, a primacy order effect still would be predicted. That is, for judgment situations involving a lengthy set of information (similar to that used in the present study), the belief-adjustment model predicts primacy.

[5] The industry information was modified from Christ (1988).

materials or to confer with one another while completing the case. No violations of these control procedures were observed during the administration of the experiment. On average, subjects required about 60 minutes to complete experiment 1.

Design and Procedure

To test for the effects of explanation on audit judgments, a 2×1 design is used. The independent variable is whether the explanation for the hypothesized non-error cause is written *prior to* or *after* the elicitation of an initial probability for the non-error cause and is manipulated between subjects. The dependent variable is the subject's assessment of the probability that the non-error cause accounts for substantially all of the gross margin fluctuation. "Substantially all" was defined in the case as a specific dollar amount (approximately 90 percent) of the gross margin increase.

Figure 1 depicts the procedure for experiment 1. Initially, the subjects were given parts A and B of the case materials and were told to read and think about the case materials. Subsequent to an initial computation task (in part C) designed to familiarize the subjects with the financial information, all subjects were told to focus on an unexpected increase in the gross margin from the prior year (i.e., the unexpected fluctuation). After learning the client's suggested cause (i.e., change in sales mix) for the gross margin increase, subjects assigned to the *treatment condition* were asked to develop a written explanation documenting how the client's cause could have accounted for substantially all of the gross margin increase. Importantly, the subjects were told that their explanation should be regarded as a "pre-workpaper explanation." This instruction was intended to avoid potential confounding from justification effects (cf. Gibbins 1984).

After subjects completed their explanation, they were instructed to assess the probability that the hypothesized non-error cause suggested by the client accounted for substantially all of the gross margin increase. This probability assessment question was stated as follows:

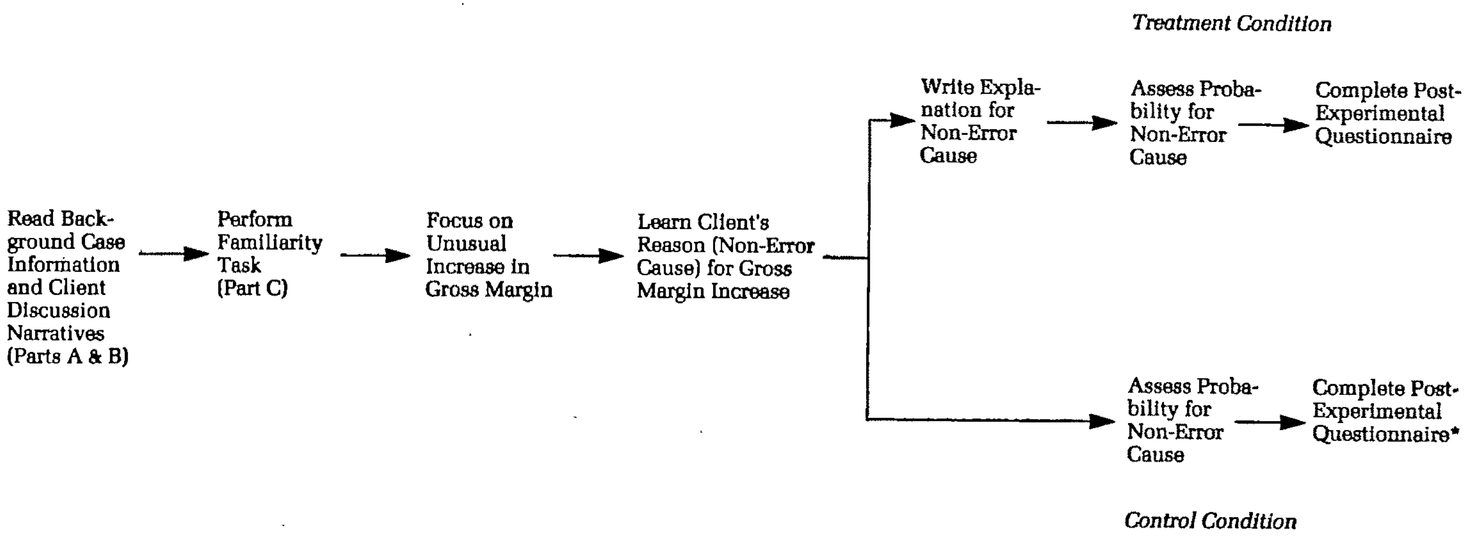
> Consider [the client's] statement about the gross margin increase. Based on your assessment of the evidence collected so far, please assess your judgment of the likelihood that a change in sales mix accounts for *substantially all* of the gross margin increase. That is, how likely do you think it is that a change in sales mix (alone) accounts for *at least $3,300,000* of the $3,667,250 gross margin increase?

After learning the client's suggested cause for the gross margin increase, subjects assigned to the *control* condition were asked to make a probability judgment about the non-error cause (i.e., they were asked to respond to the probability-assessment question presented above) *before* being asked to develop a written explanation for that cause. Next, the control-condition subjects were asked to provide a written explanation and to make a second, post-explanation probability assessment. The second probability question was similar to the initial probability question and offered a *supplementary* within-subjects test of the explanation effect.

Preliminary Analysis

Preliminary data analysis included a search for possible firm effects, text-order effects, and written responses clearly indicating an inability to successfully explain the non-error cause. No firm or text-order effects were noted for any of the probability measures (smallest p-value = 0.1353, two-tailed F-tests). Of the 34 written explanations

Figure 1
Procedure for Experiment One


Treatment Condition
Read Background Case Information and Client Discussion Narratives (Parts A & B)
Perform Familiarity Task (Part C)
Focus on Unusual Increase in Gross Margin
Learn Client's Reason (Non-Error Cause) for Gross Margin Increase
Write Explanation for Non-Error Cause
Assess Probability for Non-Error Cause
Complete Post-Experimental Questionnaire
Assess Probability for Non-Error Cause
Complete Post-Experimental Questionnaire*
Control Condition

---

* To equate the time spent on the task between the control and treatment groups and to provide a supplemental, within-subjects test of the explanation effect, the control-condition subjects also were asked to write an explanation for the non-error cause and to reassess the non-error probability. The development of the written explanation and the reassessment of the non-error probability were performed after their initial probability assessment but before completion of the post-experimental questionnaire.

obtained from the subjects (17 in each condition), eight were judged to be unsuccessful attempts at explaining the non-error cause as accounting for substantially all of the gross margin increase.[6] One subject made a computational mistake and five subjects considered the wrong products, thus precluding them from explaining an amount over the judgment threshold. The remaining two subjects made logic errors that caused them not to believe that the change in sales mix could account for substantially all of the gross margin increase. The following analysis, then, is based on data from 26 subjects (13 subjects each in the two experimental conditions).[7]

Test of Hypothesis H1

Two t-tests were performed. The first involved a between-subjects *t*-test to compare the initial, pre-explanation probability provided by the control-condition subjects with the post-explanation probability from the treatment-condition subjects. The results indicate that the final probability assessment of 76.62 provided by the treatment-condition subjects was significantly greater than the initial, pre-explanation probability of 64.62 provided by the control-condition subjects ($t=2.44$, $p=0.0112$). Thus, the null version of hypothesis H1 can be rejected in favor of the alternative, which predicted that auditors would assess higher probabilities for the hypothesized non-error cause after they had developed a written explanation for that cause.

The second *t*-test involved a within-subjects test using only the control-condition responses. The results of this paired *t*-test indicated that the final probability assessment of 75.62 by the control-condition subjects was significantly greater than their initial, pre-explanation probability of 64.62 ($t=3.26$, $p=0.0035$).[8] Thus, like the between-subjects analysis, this within-subjects test also showed a significant effect for explanation.

Consistent with psychological research but inconsistent with prior auditing research, instructing the auditor-subjects to develop a written explanation for the non-error cause led them to hold stronger beliefs for that cause, relative to the beliefs of auditors who had not yet been asked to develop a written explanation. This result suggests that auditors who write explanations are more prone to accepting an incorrect cause than auditors who do not write such explanations.

III. Experiment 2—The Counterexplanation Effect and the Primacy Order Effect

Materials and Subjects

The case materials used for the second experiment were similar to those used for the first. The primary differences pertained to the client-discussion narratives and case questions. Unlike experiment 1, the client-discussion narratives for experiment 2 contained information pertaining to other potential causes of the gross margin increase (primarily financial-statement errors) as well as information pertaining to conditions that could cancel or negate the effects of the change in sales mix. That is, they contained potential counterexplanation information. Accordingly, questions pertaining to counterexplanation were incorporated into the second (but not the first) experiment's case materials.

---

[6] The criteria used for judging a successful explanation were twofold. First, if the subject's explanation identified products, quantities, or selling prices, correct identification of such factors was required. Second, if the subject attempted to quantify the magnitude of the change in sales mix, these calculations were required to be free of error—both in terms of logic and mathematical accuracy.

[7] For both experiments 1 and 2, the hypothesis-test results remain the same if the excluded subjects are left in the analysis.

[8] The standard deviations for the pre- and post-explanation probabilities in the control condition were 12.21 and 14.90, respectively. The standard deviation for the post-explanation probabilities in the treatment condition was 12.86.

Participants in the second experiment were 53 senior-level auditors from five public accounting firms. These auditors had an average of 3.2 years of audit experience with a range of 1.9 to 5.3 years. The second experiment was conducted concurrently with the first and followed similar administration procedures. On average, subjects required about 75 minutes to complete experiment 2.

Design and Procedure

A one-way design is employed for this experiment, for which the independent variable—type of explanation procedure—is manipulated between subjects. The three levels of this independent variable are: explanation only, explanation then counterexplanation, and counterexplanation then explanation. The dependent variable is the same as that used for the first experiment—the subject's assessment of the probability that the non-error cause offered by the client accounts for substantially all of the unexpected fluctuation.

Figure 2 depicts the procedure for experiment 2. Subjects were initially asked to read the background information in part A of the case materials. Next, subjects were instructed to move ahead to part C of the case materials, which contained the case questions. Subsequent to an initial computation task, all subjects were told to focus on an unexpected increase in the gross margin from the prior year (i.e., the unexpected fluctuation) and were told the client's suggested cause for the fluctuation—a change in sales mix. Next, subjects were given written narratives documenting their prior "discussions" with various client personnel (i.e., part B of the case materials).[9]

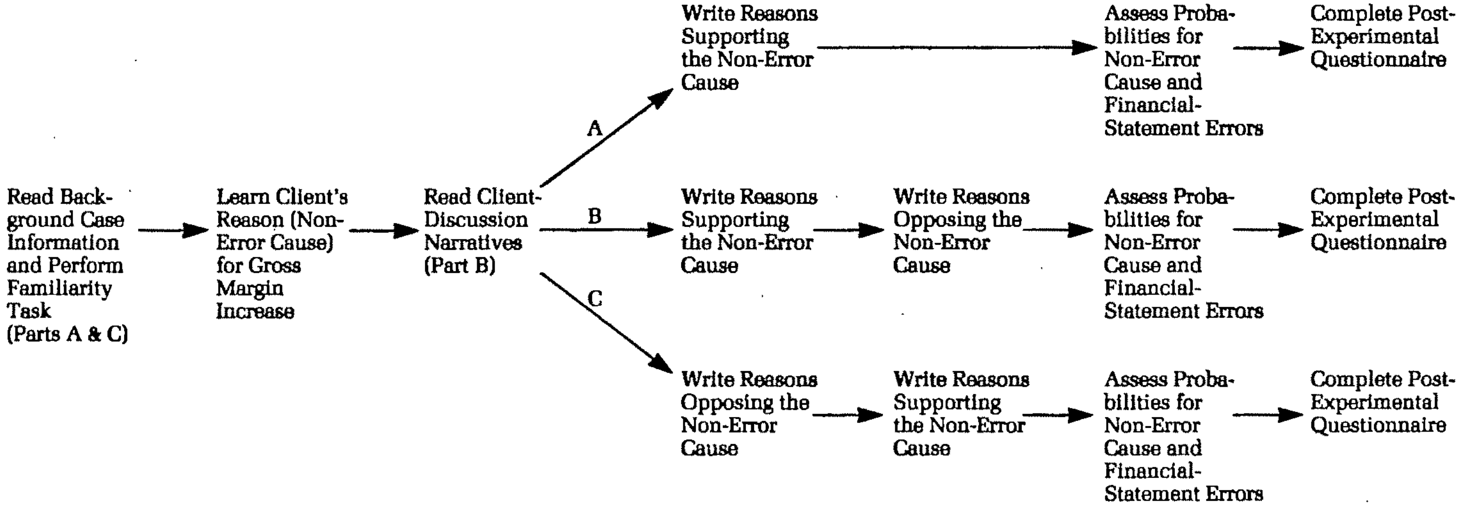
After reading the discussion narratives, subjects assigned to the *explanation-only* (E) condition were asked to document arguments supporting the non-error cause (change in sales mix) as producing substantially all of the gross margin increase. Once the change-in-sales-mix explanation was written, these subjects were asked to assess the probability that this cause accounted for substantially all of the gross margin increase. The probability assessment question was identical to that from experiment 1. As a check on the perceived multiplicity of causation, subjects were subsequently asked to assess the probability that a material financial-statement error could have resulted because of each of seven types of errors. Because of the potential for multiple causes (Shaklee and Fischhoff 1982), the subjects were told that their non-error and error probabilities need not sum to 1.00.

After reading the client-discussion narratives, subjects assigned to the *explanation-then-counterexplanation* condition (E/CE) were asked to document arguments supporting the change in sales mix as causing substantially all of the gross margin increase.

---

[9] The time at which the subjects read the client-discussion narratives differed between experiments 1 and 2. In the first (second) experiment, subjects read the narratives before (after) they learned the client's suggested cause. Since the objective for experiment 2 was not whether auditors spontaneously generate causal explanations (as it was for experiment 1), the requirement that subjects have read all available information before responding to any of the case questions was not necessary for experiment 2.

Figure 2
Procedure for Experiment Two


Read Back-ground Case Information and Perform Familiarity Task (Parts A & C)
Learn Client's Reason (Non-Error Cause) for Gross Margin Increase
Read Client-Discussion Narratives (Part B)
A
B
C
Write Reasons Supporting the Non-Error Cause
Assess Proba-bilities for Non-Error Cause and Financial-Statement Errors
Complete Post-Experimental Questionnaire
Write Reasons Supporting the Non-Error Cause
Write Reasons Opposing the Non-Error Cause
Assess Proba-bilities for Non-Error Cause and Financial-Statement Errors
Complete Post-Experimental Questionnaire
Write Reasons Opposing the Non-Error Cause
Write Reasons Supporting the Non-Error Cause
Assess Proba-bilities for Non-Error Cause and Financial-Statement Errors
Complete Post-Experimental Questionnaire

A=Explanation-only condition
B=Explanation-then-counterexplanation condition
C=Counterexplanation-then-explanation condition

After completing this task, subjects were instructed to document arguments that did *not* support this non-error cause. The latter arguments included evidence about alternative potential causes (error or non-error) of the gross margin increase as well as evidence pertaining to conditions that could cancel or negate the effects of the change in sales mix. After the explanation and counterexplanation tasks, subjects assessed the non-error cause probability and the seven financial-statement error probabilities. Subjects assigned to the *counterexplanation-then-explanation* condition (CE/E) followed essentially the same procedure, except for the reversal of the explanation and counterexplanation tasks.

Preliminary Analysis

Preliminary data analysis included a search for possible firm effects, text-order effects, and written responses clearly indicating an inability to successfully explain the non-error cause. No firm or text-order effects were noted for any of the probability measures (smallest p-value$=0.6347$, two-tailed F-tests). With the same criteria employed in experiment 1 for judging a successful explanation, seven out of 53 responses were judged to be unsuccessful attempts at explaining the change in sales mix as accounting for substantially all of the gross margin increase. Of these seven subjects, three did not identify any components to the change in sales mix nor did they perform calculations. Their written comments clearly indicated that they did not believe that the change in sales mix could account for substantially all of the gross margin increase. Two subjects excluded products from the change-in-sales-mix computation and, thus, were unable to explain an amount over the judgment threshold. The usage of an incorrect product price and a computational error precluded the remaining two subjects from successfully explaining the change in sales mix. After excluding these seven from the sample, 46 subjects remained, 15 each in the E and CE/E conditions and 16 in the E/CE condition.

Tests of Hypotheses H2 and H3

Table 1 presents both the descriptive statistics for the non-error probabilities assessed by auditors in the three experimental conditions and the planned comparisons for purposes of testing hypotheses H2 and H3. The planned comparison for hypothesis H2 shows that the mean probability assessment from auditors in the E condition (77.20) was significantly higher than the mean probability assessment from those in the counterexplanation conditions (59.65) ($t=3.16$, $p=0.0015$).[10] Thus, the null version of hypothesis H2 can be rejected in favor of the alternative, which suggests that auditors' probability judgments will be reduced once they are explicitly requested to counterexplain. This result suggests that the beliefs of auditors who have written only an explanation for the non-error cause are more likely to persevere in the presence of potentially disconfirmatory information than are the beliefs of auditors who also have written a counterexplanation. Further analysis of this result is presented in the next subsection.

Table 1 also presents the descriptive and inferential statistics for the non-error

---

[10] Supplemental analysis indicates that subjects in both counterexplanation conditions had significantly lower non-error probabilities than the explanation-only subjects. That is, E/CE subjects assessed a significantly lower mean probability (53.25) than did E subjects (77.20) ($t=4.33$, $p=0.0001$), and CE/E subjects also assessed a significantly lower mean probability (66.47) than did the E subjects (77.20) ($t=1.78$, $p=0.0449$).

Table 1
Descriptive Statistics and Planned-Comparison Tests for the Non-Error Probabilities—Experiment 2

Panel A. Descriptive Statistics:

	(A) Explanation Only (n=15)	(B) Explanation Then Counterexplanation (n=16)	(C) Counterexplanation Then Explanation (n=15)	(D) Columns B & C Averaged (n=31)
Mean Probability Judgment	77.20	53.25	66.47	59.65
Standard Deviation	10.93	19.01	20.67	20.62

Panel B. Planned-Comparison Tests:

Planned Comparison	*Mean Square*	*t-statistic*	*Probability (One-tailed)*
Column A vs. D	3038.99	3.16	0.0015
Column B vs. C	1352.36	−2.11	0.9796

probabilities assessed by the auditors in the two counterexplanation conditions—the two conditions used to test for the primacy order effect predicted by hypothesis H3. A comparison of the sample means indicates that subjects assigned to the E/CE condition assessed lower probabilities for the non-error cause (53.25) than did subjects assigned to the CE/E condition (66.47). These sample means are consistent with a recency order effect, rather than the hypothesized primacy order effect. Thus, the null version of hypothesis H3 cannot be rejected in favor of the alternative, which predicted that a primacy order effect should occur ($t=-2.11$, $p=0.9796$). Further analysis of this result also is presented in the next section.

Other Analyses

Other data collected as part of the experiment are useful in explaining the second experiment's results. Recall that subjects in all three experimental conditions assessed the probability of occurrence for seven financial-statement errors. Table 2 presents the auditors' assessments of financial-statement error—averaged across the seven error types.[11] As expected, the counterexplanation-condition subjects (separately or combined) considered the occurrence of a financial-statement error to be more probable than did the explanation-only subjects (all p-values<0.05). That is, the likelihood of financial-statement error was greater when explicit counterexplanation occurred.

To rule out the possibility that the explanation-only subjects' non-error (average error) probability assessments were systematically higher (lower) than those rendered

---

[11] Although the mean probabilities for each error type could be analyzed separately, this approach is problematic since there is significant variability in the number of subjects explaining each error type. Thus, to provide a less variable measure, the average error measure (across the seven financial-statement errors) was used.

Table 2
Descriptive Statistics and Planned-Comparison Tests for the Average Financial-Statement Error Probabilities—Experiment 2

Panel A. Descriptive Statistics:

	(A) *Explanation Only* (n=15)	(B) *Explanation Then Counterexplanation* (n=16)	(C) *Counterexplanation Then Explanation* (n=15)	(D) *Columns B & C Averaged* (n=31)
Mean Probability Judgment	35.28	49.12	45.47	47.35
Standard Deviation	12.80	10.39	8.98	9.75

Panel B. Planned-Comparison Tests:

Planned Comparison	*Mean Square*	*F-statistic*	*Probability (Two-tailed)*
Column A vs. D	1458.84	12.45	0.0010
Column A vs. B	1482.91	12.65	0.0009
Column A vs. C	778.84	6.64	0.0135
Column B vs. C	103.11	0.88	0.3536

by the counterexplanation subjects because of more elaborate explanations provided by the former subjects, a content analysis was performed to determine the mean number of explanation arguments provided by each subject.[12] The content analysis revealed no significant difference in the total number of explanation arguments generated between the explanation and the two counterexplanation conditions combined ($F=1.34$, $p=0.2530$).

When comparing the average error probability between the two counterexplanation conditions, no significant difference in probability emerges ($F=0.88$, $p=0.3536$). The CE/E subjects were expected to assess a larger error probability than the E/CE subjects since the theoretical development for the primacy order effect suggested that the multiplicity of causation would be greater when counterexplanation occurs initially (CE/E) rather than after explanation (E/CE). Even though there was no difference in the financial-statement error probabilities between the two counterexplanation conditions, post-experimental data reveal that the E/CE subjects perceived the counterexplanation evidence to be stronger than did the CE/E subjects ($F=4.17$, $p=0.0472$). This result, in conjunction with the finding that the E/CE subjects judged the explanation evidence to be weaker than did the CE/E subjects ($F=5.53$, $p=0.0234$), is consistent with the experimental results for the non-error probability data—that is, consistent with a recency order effect. The evidence considered last (first) in the two-task sequence was deemed to be stronger (weaker) than if it were considered first (last).

---

[12] The content analysis was performed independently by the researcher and a paid coder. Intraclass correlations for agreement on counts of total arguments were separately computed for the explanation and, where applicable, counterexplanation responses. All statistics indicated statistically significant agreement between the two coders ($p<0.0001$). Any differences in coding were jointly resolved.

A possible explanatory variable for the unexpected recency results pertains to the subjects' perceptions of task difficulty. Post-experimental data indicate that the CE/E subjects found counterexplanation significantly more difficult than did the E/CE subjects ($F=6.13$, $p=0.0173$).[13] Since prior research indicates that there is an inverse relation between task difficulty and importance weights assigned to evidence in the judgment-formulation process (Sherman et al. 1985), these results suggest that counterexplanation evidence has a smaller impact on judgments about the non-error cause when generated initially rather than after explanation. Indeed, these results may indicate that auditors must comprehend (i.e., explain) an hypothesized cause before they can disconfirm it via counterexplanation (cf. Gilbert 1991).

Finally, to rule out the possibility that task difficulty affected the number of causal arguments comprising the explanations and counterexplanations, several additional content-analysis comparisons were made. The comparisons revealed no significant difference in the total number of counterexplanation arguments generated by the two counterexplanation conditions ($F=1.83$, $p=0.1833$). The content analysis did, however, indicate a significant difference in the total number of explanation arguments. Specifically, the CE/E subjects assessed a smaller number of explanation arguments (8.74 on average) than did E/CE subjects (13.44 on average) ($F=6.47$, $p=0.0146$). A more detailed analysis of the written explanations reveals that many of the CE/E subjects documented no more than a change-in-sales-mix computation for their written explanation. Most E/CE subjects, in contrast, elaborated to a greater extent by also documenting other supporting arguments from the case materials. This content analysis result is consistent with research in which primacy, rather than recency, order effects are observed (Hoch 1984; Moser 1989). Since the non-error sample means from the two counterexplanation conditions are in the direction of a recency order effect, it must be concluded that differential weighting, rather than differential argument production, produced the recency results.

IV. Concluding Comments

Experiment 1 demonstrates that auditors asked to develop and document explanations supporting the validity of a non-error cause consider its likelihood of occurrence to be significantly greater than do auditors not asked to develop and document such explanations. As shown by experiment 2, the explanation-effect result may have significant potential implications for audit effectiveness. That is, once an auditor has developed a written explanation for a non-error cause, belief in the validity of that cause appears to persevere unless the auditor is explicitly requested to counterexplain. Hence, auditors do not spontaneously counterexplain and, thus, reduce their beliefs in the inherited non-error cause to the same extent as when written counterexplanation occurs. Since current auditing standards do not require auditors to engage in counterexplanation when performing analytical review, these results suggest that auditors who do not explicitly consider why an inherited non-error cause may *not* be correct (and, thus, consider what error causes may be correct) will be more prone to compromising audit effectiveness by accepting a potentially incorrect non-error cause. By identifying an analytical review situation in which a compromise of audit effectiveness may occur, this study extends other research by Heiman (1990) and Bedard and Biggs (1991a, 1991b), which investigated situations in which compromises of audit efficiency were likely (i.e., situations in which the auditor was prone to accepting the wrong error cause).[14]

---

[13] Although the E/CE subjects found it more difficult to generate their written explanations than the CE/E subjects, this difference was not statistically significant ($F=0.80$, $p=0.3750$).

Other results from experiment 2 indicate that a primacy order effect does not occur when explanation and counterexplanation activities are sequentially performed in the two possible task-performance orders. Rather, the experimental data are in the direction of a recency order effect. Since post-experimental data indicate that CE/E auditors had greater difficulty generating opposing arguments than did E/CE auditors, it may be that auditors must comprehend (via explanation) a non-error cause before they can disconfirm it (cf. Gilbert 1991). That is, the natural order of comprehension may be to explain and then (possibly) counterexplain.

Although the present study's results may have implications for how audit firms structure analytical review, additional research is necessary before any changes in practice are in order. Such research should determine whether other distinctive features of the audit context (which were not investigated herein) mitigate the potential auditor biases identified in this study. For example, auditor accountability (cf. Tetlock 1983) and the audit review process (Solomon 1987) may counteract any potentially biasing effects associated with auditors developing written explanations. These two factors also may increase an auditor's predisposition to counterexplain spontaneously, thereby reducing the likelihood of acceptance of an incorrect cause. Given the increased reliance placed on audit evidence obtained through analytical review, understanding how these (and other) distinctive features of the audit environment influence the effectiveness of analytical review is a promising direction for future research.

[14] In these situations, the auditor expends unnecessary audit effort investigating the wrong error cause, thereby compromising audit efficiency. Since the auditor probably would not cease the investigation until the "correct" error cause was identified, it is unlikely that audit effectiveness would be compromised in these situations.

References

American Institute of Certified Public Accountants. 1982. *Working Papers: Statement on Auditing Standards No. 41*. New York: AICPA.

———. 1988a. *Analytical Procedures: Statement on Auditing Standards No. 56*. New York: AICPA.

———. 1988b. *The Auditor's Responsibility to Detect and Report Errors and Irregularities: Statement on Auditing Standards No. 53*. New York: AICPA.

Anderson, C. 1982. Inoculation and counterexplanation: Debiasing techniques in the perseverance of social theories. *Social Cognition* 1: 126–39.

———, M. Lepper, and L. Ross. 1980. Perseverance of social theories: The role of explanation in the persistence of discredited information. *Journal of Personality and Social Psychology* 39 (December): 1037–49.

———, B. New, and J. Speer. 1985. Argument availability as a mediator of social theory perseverance. *Social Cognition* 3 (Fall): 235–49.

———, and E. Sechler. 1986. Effects of explanation and counterexplanation on the development and use of social theories. *Journal of Personality and Social Psychology* 50 (January): 24–34.

Anderson, N. 1981. *Foundations of Information Integration Theory.* New York: Academic Press.

Anderson, U., and W. Wright. 1988. Expertise and the explanation effect. *Organizational Behavior and Human Decision Processes* 42 (October): 250–69.

Asare, S. 1991. The auditor's going-concern decision: Interaction of task variables and the sequential processing of evidence. Working paper, University of Florida, Gainesville (June).

Ashton, A., and R. Ashton. 1988. Sequential belief revision in auditing. *The Accounting Review* 63 (October): 623–41.

Bedard, J., and S. Biggs. 1991a. The effect of domain-specific experience on evaluation of management representations in analytical procedures. *Auditing: A Journal of Practice & Theory* 10 (Supplement). Forthcoming.

———, and ———. 1991b. Pattern recognition, hypotheses generation, and auditor performance in an analytical task. *The Accounting Review* 66 (July): 622–42.

Christ, M. 1988. Attention and encoding during audit planning: An experimental study using a schematic framework. Unpublished dissertation, University of Texas, Austin.

Einhorn, H., and R. Hogarth. 1986. Judging probable cause. *Psychological Review* 99 (January): 3–19.

Gibbins, M. 1984. Propositions about the psychology of professional judgment in public accounting. *Journal of Accounting Research* 22 (Spring): 103–25.

Gilbert, D. 1991. How mental systems believe. *American Psychologist* 46 (February): 107–19.

Heiman, V. 1990. Auditors' assessments of the strength of the likelihood of error explanations in analytical review. *The Accounting Review* 65 (October): 875–90.

Hoch, S. 1984. Availability and interference in predictive judgment. *Journal of Experimental Psychology: Learning, Memory and Cognition* 10 (October): 649–62.

Hogarth, R., and H. Einhorn. Order effects in belief updating: The belief-adjustment model. *Cognitive Psychology* (forthcoming).

Kinney, W. 1987. Attention directing analytical review using accounting ratios: A case study. *Auditing: A Journal of Practice & Theory* 6 (Spring): 59–73.

Kreutzfeldt, R., and W. Wallace. 1986. Error characteristics in audit populations: Their profile and relationship to environmental factors. *Auditing: A Journal of Practice & Theory* 6 (Fall): 20–43.

Libby, R. 1985. Availability and the generation of hypotheses in analytical review. *Journal of Accounting Research* 23 (Autumn): 648–67.

———, and D. Frederick. 1990. Experience and the ability to explain audit findings. *Journal of Accounting Research* 28 (Autumn): 346–67.

Moser, D. 1989. The effects of output interference, availability, and accounting information on investors' predictive judgments. *The Accounting Review* 64 (July): 433–48.

Nisbett, R., and L. Ross. 1980. *Human Inference: Strategies and Shortcomings of Social Judgment.* Englewood Cliffs, NJ: Prentice-Hall.

Pei, K., C. Moeckel, and H. Reneau. 1991. An examination of the relationship between auditor judgment and memory for supporting evidence. Working paper, Arizona State University, Tempe (April).

Shaklee, H., and B. Fischhoff. 1982. Strategies of information search in causal analysis. *Memory and Cognition* 10 (November): 520–30.

Sherman, S., R. Cialdini, D. Schwartzman, and K. Reynolds. 1985. Imagining can heighten or lower the perceived likelihood of contracting a disease: The mediating effect of ease of imagery. *Personality and Social Psychology Bulletin* 11 (March): 118–27.

Solomon, I. 1987. Multi-auditor judgment/decision making research. *Journal of Accounting Literature* 6: 1–25.

Tetlock, P. 1983. Accountability and the perseverance of first impressions. *Social Psychology Quarterly* 46 (December): 285–92.

Tubbs, R., W. Messier, and R. Knechel. 1990. Recency effects in the auditor's belief-revision process. *The Accounting Review* 65 (April): 452–60.

Wright, A., and R. Ashton. 1989. Identifying audit adjustments with attention-directing procedures. *The Accounting Review* 64 (October): 710–19.

THE ACCOUNTING REVIEW
Vol. 67, No. 1
January 1992
pp. 77–95

The Effect of Antitrust Investigations on Discretionary Accruals: A Refined Test of the Political-Cost Hypothesis

Steven F. Cahan
University of Wyoming

SYNOPSIS AND INTRODUCTION: The antitrust laws of the United States prohibit monopolies or attempts to create a monopoly in any unregulated line of business. In the past, the two agencies that enforce these laws, the Department of Justice and the Federal Trade Commission, have relied on accounting profits in prosecuting such antitrust violations. The agencies argued that high accounting rates of return were "excessive" and indicative of monopolistic power on the part of the firm. Thus, to reduce the possibility of an unfavorable ruling and the costs associated with it, managers in firms investigated for monopoly-related violations would have incentive to use accounting procedures (e.g., accounting methods, accruals) that produce abnormally low levels of income. Because the incentive to reduce income will increase as the threat of an unfavorable ruling becomes more imminent, it is expected that managers will take additional steps to lower income while being actively investigated, compared with periods of non-investigation.

The differences in political costs caused by the antitrust investigation are used to provide a refined test of the political-cost hypothesis. In particular, the study examines, on a longitudinal basis, whether managers respond to these investigations by adjusting their discretionary accruals. The discretionary accruals for 48 firms that were investigated for monopoly-related violations were estimated over a 15-year period by using the residuals of a fixed effects covariance model that regressed total accruals on the change in sales, the fixed asset balance, and dummy variables

I am very grateful to the editor and two anonymous referees for their helpful comments. I also thank Susan Moyer, James Manegold, Frank Selto, Phil Shane, Dennis Schmidt, and workshop participants at the University of Wyoming for reading and commenting on earlier drafts of this article. This research was funded by a University of Wyoming summer research grant.

Submitted September 1989.
Accepted June 1991.

representing each firm and year. The specific hypothesis tested is whether the estimated discretionary accruals over the period of investigation (typically more than one year) were lower, or more income-reducing, than those in other years.

The hypothesis was tested by using a nested design with a dummy variable, coded 1 for the years of investigation, included in the accrual model. This variable was significant and negatively signed, which indicates that discretionary accruals were lower while the firm was being investigated, as expected. The discretionary accruals for a control group of firms that were not investigated were also estimated, and these were found to remain the same during the periods of investigation and non-investigation for the matched sample firms. The results from these tests support the political-cost hypothesis and are consistent with the view that managers adjust earnings in response to monopoly-related antitrust investigations.

Key Words: *Political costs, Discretionary accruals, Antitrust.*

Data Availability: *A hard copy of the data can be obtained from the author on request.*

THE remainder of this article is divided into five sections: Section I provides background on the political-cost hypothesis. Section II develops the research hypothesis. Section III describes the data and method. Section IV reports results, and section V is a conclusion.

I. Background

The political-cost hypothesis predicts that managers confronted with the possibility of politically imposed wealth transfers will choose accounting strategies that reduce the likelihood or size of the transfer. Watts and Zimmerman (1978) argue that politicians and their constituents equate high accounting profits with monopolies and either lack incentive or are unable to undo accounting numbers completely. Thus, high income firms will be particularly vulnerable to wealth-extracting political transfers in the form of legislation and/or regulation, and managers in these firms will have incentive to use accounting procedures that reduce reported income.

Watts and Zimmerman (1978) and subsequent researchers generally have used firm size to measure the firm's vulnerability to political costs. The usual finding is a positive relation between firm size and the use of income-lowering accounting methods, which is interpreted as supporting the political-cost hypothesis (see Watts and Zimmerman 1986 for a review). However, evidence contrary to the political-cost hypothesis has been found (see, e.g., Bowen et al. 1981; Moyer 1990; Zimmer 1986), and some researchers (e.g., Ball and Foster 1982) argue that the size proxy needs to be more plausibly and explicitly linked to political costs. Thus, prior results are difficult to interpret because of the crudeness of the size proxy.

Recently, researchers have begun to explore alternative proxies for political costs. For example, Sutton (1988) notes that because of price controls in effect in the United Kingdom in the 1970s, U.K. firms with excessively high profit margins were likely to come under the scrutiny of government regulators. Sutton goes on to use profit margin as a proxy for political cost. Likewise, Wong (1988) reviews the tax reform movement in New Zealand in the early 1980s to support the use of such proxies as reported tax rates and export credit sales for political costs.

These alternates, while providing better mappings of the relation of political cost to the proxy, still ignore the mechanics of the political process. To be specific, political action follows a complex decision on the part of the politician (or regulator). For example, the work of Stigler (1971) and Peltzman (1976) suggests that the decisions of vote-maximizing politicians are swayed by such factors as the size of the wealth transfer, the number of public voters, and the political funds provided by the firm. Further, Watts (1977) suggests the likelihood that political action will occur is based largely on the presence of "crises." Given the importance of these other factors, it is unclear how closely proxies such as a firm's profit margin or tax rate are related to political costs.

II. Research Hypothesis

Design Issues

To provide a more rigorous test of the political-cost hypothesis, this study differs from prior research in two ways. First, the sample is composed of firms that were actual targets of a particular political action. Second, the analysis is done on a longitudinal basis.

Traditionally, the political-cost hypothesis has been tested by studying how firms or their stock prices react to hypothetical costs arising from unspecified but potential, politically imposed wealth transfers. The power of these tests relies on the adequacy of proxies, such as firm size, for identifying firms exposed to these risks. The approach taken in the present study is to examine only those firms that were actually targeted by a specific political action. The advantage is twofold: The politician's complex selection process is made exogenous, and if the political action is significant, potential political costs to affected firms are largely assured.

Prior research also has examined the accounting choices of firms on a cross-sectional basis. However, even for firms affected by a similar political action, the relative political costs are difficult to quantify. This study instead uses a longitudinal approach that exploits likely changes in a single firm's political costs over time.[1]

The manager faced with an exogenous political decision based somewhat on the firm's income has incentives to reduce reported accounting income. These incentives increase as the firm's expected political costs increase. The firm's expected political costs are a joint function of the probability that the action will be undertaken, the probability that the outcome of the action will be unfavorable, and the possible costs to the firm if the outcome is unfavorable. Thus, once action is taken, expected political costs increase because the first probability becomes one, and a manager would be expected to respond to this increase by engaging in activities that reduce income. However, when the final outcome of the action is decided, the probabilistic determination of related political costs become irrelevant, and expected political costs related to this action decrease, since the outcome and the imposed costs, if any, are known with certainty. At this point, the manager would no longer have an incentive to reduce income.

---

[1] Recently, Jones (1990) used a similar approach in looking at firms seeking political benefits (i.e., import trade protection) from the International Trade Commission.

The Antitrust Process

Monopoly-related antitrust investigations were chosen as the political action of interest for two reasons. First, the Department of Justice (DOJ) and Federal Trade Commission (FTC) have used accounting numbers in prosecuting these cases. In the late 1950s, these agencies began to rely on relatively high accounting rates of return as evidence of excessive market power (see, e.g., Meehan and Larner 1989). For example, accounting-based analysis played a substantial role in cases involving IBM, Borden, and General Mills (Elzinga 1989; Fisher et al. 1983). This is critical because managers have incentives to reduce earnings only if income is used in making antitrust decisions. Although the Reagan administration later changed this focus, during the period covered in this study (1970–83), accounting profits were important evidence in antitrust cases.

Second, the political costs associated with an unfavorable antitrust ruling can be extremely high. The Antitrust Division of the DOJ has authority to bring civil and/or criminal actions against the firm or its management. The intent of a civil case is to end the unlawful conduct and prevent its recurrence in the future. The settlement of these cases is usually by a consent decree in which possible outcomes include divestiture of stock or assets, dissolution of the offending organization, forced disclosure of certain patents or know-how, or alteration of specific terms by which goods are being bought or sold (Howard 1983). Criminal cases can lead to fines of up to $1 million for the corporation and fines of up to $100,000 and prison terms not exceeding three years for the individual. The FTC settles cases in a similar manner, using a consent order, but differs from the DOJ in using administrative rather than judicial procedures and in pursuing only civil actions.

These substantial political costs provide incentives to management to reduce income. By lowering income, the manager may reduce political costs in various ways. Possible outcomes of a DOJ or FTC investigation include dismissal of the case and voluntary or involuntary compliance with antitrust statutes. The manager could reduce expected political costs by increasing the probability of a dismissal. By reducing income during the period of investigation (and thereby making the firm's accounting rate of return more comparable to that of other firms in the industry), the manager could argue that the firm's monopoly had diminished. Even if dismissal was unlikely, the manager might point to the lower rate of return as evidence that the firm cannot afford a costly settlement and argue for a lighter penalty.

Accrual Choices

The manager has several mechanisms to lower income, including choosing accounting methods or accruals that defer income to later periods. The earliest studies of the political-cost hypothesis typically examined choices between alternative accounting methods (e.g., expensing/capitalizing R&D, successful efforts/full cost), but later studies, beginning with Healy's (1985) test of the related bonus-plan hypothesis, examine the firm's accrual choices.

Tests of accounting methods would not be appropriate in the current study. For one thing, most of a firm's accounting methods are chosen prior to the initiation of antitrust action. Although managers could choose to switch accounting methods, such a change would be visible and could be undone by the enforcing agency. Moreover, to the extent that firms are aware of their political exposure and already using income-reducing methods, the potential for further changes once an investigation begins is limited.

In contrast, accrual choices are less visible and less likely to be undone by political decision makers. The undoing of income differences caused by accruals can be difficult because the information required to make these adjustments may not be available (Schipper 1989). Further, as Scholes et al. (1989) contend, regulators rationally use simple decision models because of limited processing time and ability and so may not make these adjustments in any case.

The manager also has more opportunity to use accruals to reduce income. Unlike accounting choices, income-decreasing accruals probably will not have been exhausted prior to an investigation. If they have, when the accruals reverse and come into income, the subsequent increase in income might actually make an investigation or unfavorable settlement more likely. Also, because accrual choices are made continuously, managers can adjust or "manage" earnings to achieve some optimal level each year. Thus, managers could use accruals to lower income while being investigated and allow these to reverse in later periods.

The total accruals of a firm are composed of discretionary accruals, DA, and nondiscretionary accruals, NA, or notationally:

$$A = DA + NA, \qquad (1)$$

where A represents total accounting accruals. Both discretionary and nondiscretionary accruals are made within the constraints of generally accepted accounting principles. However, discretionary accruals allow leeway in the reported amounts, and the manager chooses an accrual within an acceptable range.

Thus, this research operationalizes the political-cost hypothesis as:

H1: The discretionary accruals of firms that were investigated for antitrust violations were more income-reducing during the investigation than in periods when the firm was not being investigated.

Discretionary Accruals

Because the two components of total accruals cannot be observed, it is necessary to develop a proxy for discretionary accruals. For example, the discretionary accrual could be estimated by taking the total accrual, defined as the change in the working capital accounts less depreciation and deferred tax expenses, and subtracting out an estimate of the nondiscretionary accrual. In conformance with McNichols and Wilson (1988), this can be written as:

$$DAP = A - NAEST, \qquad (2)$$

where DAP is the proxy for discretionary accruals and where $NAEST$ is the estimate of nondiscretionary accruals.

Recently, Jones (1990) estimated nondiscretionary accruals by regressing total accruals on two variables, the change in sales and the fixed asset balance. Letting

CHGSALES equal the first variable and *FIXASSET* the second, the basic model is shown as:

$$A_{it}=B_0+B_1 CHGSALES_{it}+B_2 FIXASSET_{it}+u_{it}, \tag{3}$$

where the amount, $B_0+B_1 CHGSALES_{it}+B_2 FIXASSET_{it}$, equals *NAEST*. Thus the discretionary accruals are by definition unexpected, defined as the prediction error. The expected sign for *CHGSALES* is positive. This is because the changes in the working capital accounts, such as accounts receivable, inventory, accounts payable, and taxes payable, vary directly with levels of sales activity. For example, accounts receivable typically increase as sales increase even when the manager does not exercise any discretion over the firm's accruals. The net effect of an increase in sales should be an increase in the working capital accounts or net income increasing accruals. The expected sign for *FIXASSET* is negative since accruals for expenses such as depreciation and deferred tax will be affected by the fixed assets in place. Firms with more fixed assets would be expected to have more periodic cost allocations related to their fixed asset accounts and, in addition, more timing differences giving rise to deferred tax.[2]

The error term, u_{it}, or difference between the total accrual and the estimate of the nondiscretionary accrual can be interpreted as the part of accruals being "managed." For example, a manager could produce lower income by increasing bad debt expense, increasing inventory write-offs, classifying more indirect manufacturing costs as period expenses rather than inventoriable costs, accelerating purchases of inventory at year-end when LIFO is used, extending the use of accelerated depreciation methods, or reducing estimated lives or salvage values for fixed assets.

The approach used here and by Jones (1990) should lead to an estimate of discretionary accruals that is less biased and less noisy than those used in earlier studies. For example, Healy (1985), Liberty and Zimmerman (1986), and DeAngelo (1986) assume that nondiscretionary accruals are stationary over time.

III. Data and Method

Sample

The sample is based on 48 firms that were investigated for monopoly-related violations by the DOJ or FTC between 1970 and 1983.[3] The sample was selected by first reviewing *The Wall Street Journal Index* under the key word "antitrust." The initial screening of the *Index* for the 15-year period yielded more than 200 articles that were possibly related to antitrust actions involving a monopoly. These articles were read to determine whether the action was related to such an investigation and, if so, to identify the particular firm or firms involved. For a firm to be included in the sample, the article had to mention the word "monopoly" explicitly. This procedure minimized subjective interpretation by the researcher. The firm also had to be listed on either the COMPUSTAT industrial or research tapes.[4] Firms affected by other types of government antitrust actions, such as antimerger or price-fixing cases, and private antitrust cases were excluded from the sample because income plays a less important role in prosecuting these violations (see, e.g., Areeda 1981). The sample firms are listed in the appendix with information about their respective cases.[5]

---

[2] The change in sales and the fixed asset balance, as Jones (1990) recognizes, may not be completely exogenous. For example, managers may delay or accelerate deliveries to produce a more appropriate sales figure. Also, the deferred tax expense is not actually a "free move" but is a function of other accruals. For example, a firm using accelerated depreciation, an income-decreasing accrual, will have a small deferred tax expense. Nonetheless, because the net effect on income is jointly determined, deferred tax should be included as part of accruals.

[3] The 1970 cutoff was required because the data on the 1988 COMPUSTAT industrial and research tapes begin in 1968; because accruals cannot be computed without a lagged year, the earliest accruals are computed for 1969, and a 1970 cutoff ensures at least one investigation-free year prior to the period considered. The 1983 cutoff was required because, starting in 1984, the use of the profits measure was less widespread.

Method

Financial data items were collected from COMPUSTAT. The total accruals, A, were computed as follows (Healy 1985; Liberty and Zimmerman 1986):

$$A_t = -DEP_t + (AR_t - AR_{t-1}) + (INV_t - INV_{t-1}) - (AP_t - AP_{t-1}) - (TP_t - TP_{t-1}) - DEF_t, \tag{4}$$

where DEP is depreciation, AR is the accounts receivable balance, INV is the inventory balance, AP is the accounts payable balance, TP is the taxes payable balance, and DEF is deferred tax expense.

The total accruals in year t are computed for each firm i over the 15-year period 1969–83. The data were pooled over time and firms, and a covariance model was used (Pindyck and Rubinfeld 1981). This model assumes the residual, e_{it}, is composed of a time effect, which will be the same for all firms in a particular year, and a firm effect, which will be the same for each firm over all years.

The use of a covariance model has a number of advantages (Judge et al. 1985). First, with pooled data, a hypothesis regarding interfirm differences over time can be tested by using a single significance test as opposed to a series of individual cross-sectional tests. Second, to the extent that residual correlations are caused by fixed effects, the covariance model will reduce these correlations and yield more efficient estimates and unbiased standard errors. Third, the slope coefficients will also be less biased by omitted variables because some of the effect of these variables will be captured by the fixed effects.

The estimated discretionary accrual is defined as the residual, e_{it}, from the subsequent model:

$$A_{it}/TA_{it} = b_0[1/TA_{it}] + b_1[CHGSALES_{it}/TA_{it}] + b_2[FIXASSET_{it}/TA_{it}] + b_3 YR_{i1} + \ldots + b_{17} YR_{i13} + b_{18} FIRM_{1t} + \ldots + b_{64} FIRM_{47t} + e_{it}, \tag{5}$$

where $CHGSALES$ is the change in sales from year $t-1$ to year t and $FIXASSET$ is the balance of property, plant, and equipment at the end of year t. YR_{it} is dummy-coded as 1 for year t ($t=1969, 1983$), and $FIRM_{it}$ is dummy-coded as 1 for firm i ($i=1, 47$). The YR variables measure the time effect for each of the 15 years. The $FIRM$ variables measure the firm effect for 47 of the sample firms. These two variables, when combined, define a unique intercept for each firm/year observation. The variables, *A*, *CHGSALES*, *FIXASSET*, and the constant term, are deflated by total assets, *TA*. This is to reduce the effect of heteroscedasticity, or unequal error variances, which can lead to inefficient least-squares regression parameter estimates and biased estimates of the variances of those parameters.[6]

---

[4] *The Wall Street Journal Index* appears to be relatively complete in its coverage of monopoly-related antitrust actions. The author also examined Commerce Clearing House's *Trade Cases Reporter*, which reports on all court decisions arising from DOJ and FTC actions; however, unlike the *Index*, it does not report on investigations if no formal action was taken. This cross-check did not identify any cases that were not reported in the *Index* and that affected firms available on COMPUSTAT.

[5] A number of the industry-related investigations were brought under the "shared monopoly" doctrine. Under this viewpoint, an industry is not controlled by a single firm but rather a few firms (an oligopoly) whose behavior and economic performance is like a monopolist. See Areeda (1981) for details.

To test the political-cost hypothesis, equation (5) is expanded to include an additional variable that separates observations on the basis of whether they occurred during an investigation. The variable, *AT*, is dummy-coded as 1 for observation *it* if firm *i* was being investigated for a monopolistic violation in year *t*, and as 0 in other cases.[7] The expanded equation is written as:

$$\begin{aligned} A_{it}/TA_{it} = {} & b_0[1/TA_{it}] + b_1[CHGSALES_{it}/TA_{it}] + b_2[FIXASSET_{it}/TA_{it}] + b_3YR_{i1} \\ & + \ldots + b_{17}YR_{i13} + b_{18}FIRM_{1t} + \ldots + b_{64}FIRM_{47t} + b_{65}AT_{it}, \end{aligned} \tag{6}$$

where *AT* tests whether the presence of the investigation has incremental ability to explain a significant portion of the remaining cross-sectional, intertemporal variation in total accruals or, simply, whether the estimated discretionary accruals are correlated with the period of investigation. With this coding scheme, if managers lower accruals when under investigation, the coefficient for *AT* will be negative.

IV. Empirical Tests and Results

Table 1 reports the outcome for the regression of equation (5). The F-statistic for the model is significant at the 0.0001 level and the adjusted R^2 is 63.7 percent. The variables, *CHGSALES/TA* and *FIXASSET/TA*, were both related to total accruals at the 0.01 probability level or better. The signs of *CHGSALES/TA* and *FIXASSET/TA* are as expected. The *YR* and *FIRM* variables were tested jointly with an F-test. The F-statistic was significant at the 0.0001 level, which indicates that the dummy variables as a whole were significantly related to total accruals. The individual *YR* and *FIRM* variables are not easily interpreted. They are meaningful only when combined to form intercepts for each of the 667 firm/year observations.

The results for equation (5) support the view that the variation in accruals can be largely explained by the change in sales, the fixed asset balance, and time- and firm-dependent fixed effects. Thus, the remaining cross-sectional, intertemporal variation in accruals (the residual) is assumed to be the outcome of managerial discretion.

The residuals were examined for heteroscedasticity and autocorrelation. The mean residual is zero (s.d. = 0.035), which is expected on the basis of the mechanics of regression. The correlation between the squared residuals and squared total assets is −0.0615, which is not significant at the 0.05 level. Thus, the scaling appears to have corrected for heteroscedasticity caused by differences in firm size. The Durbin-Watson statistic is also examined since time-series data can lead to positive or negative serial

---

[6] The correlation between squared total assets and the squared residuals from an unscaled model was 0.2579, which is significant at the 0.0001 level. This indicates that the error variances are related to firm size and that deflated variables should be used.

[7] The period of investigation is defined as the year the investigation began through the year prior to its end (including settlement or dismissal of a lawsuit, if applicable). That is, if an investigation begins in 1978 and ends in 1980, the manager would be expected to have incentive to take income-reducing discretionary accruals in 1978 and 1979.

Table 1

Estimation Results for Equation (5) to Test the Effect of Change in Net Sales, Fixed Asset Balance, and Year, Firm Variables on Total Accruals

$$A_{it}/TA_{it}=b_0[1/TA_{it}]+b_1[CHGSALES_{it}/TA_{it}]+b_2[FIXASSET_{it}/TA_{it}]$$
$$+b_3YR_{i1}+\ldots+b_{17}YR_{i13}+b_{18}FIRM_{1t}+\ldots+b_{64}FIRM_{47t}$$

Independent Variable[a]	*Expected Sign*	*Coefficient*	t-*value*
CHGSALES/TA	+	0.0306	2.469[b]
FIXASSET/TA	−	−0.1020	−4.757[c]
1/TA		−0.1653	−0.336

Note: F-statistic = 18.968 (significant at 0.0001 with two-tailed test); adjusted R^2=63.65 percent; Durbin-Watson statistic = 2.020.

[a] Regression variables:

$$A=-DEP_{it}+(AR_{it}-AR_{it-1})+(INV_{it}-INV_{it-1})-(AP_{it}-AP_{it-1})-(TP_{it}-TP_{it-1})-DEF_{it},$$

where:

DEP = Depreciation;
AR = Accounts receivable;
INV = Inventory;
AP = Accounts payable;
TP = Taxes payable;
DEF = Deferred tax expense; and
$CHGSALES$ = Net sales_{it} − Net sales_{it-1};
$FIXASSET$ = Fixed assets_{it};
TA = Total assets_{it};
YR = 1 if year t (t=1, 13), 0 otherwise; and
$FIRM$ = 1 if firm i (i=1, 47), 0 otherwise.

The model was estimated by using least squares regression for 667 observations related to 48 firms for the years 1969–83. Coefficients and t-values for the dummy variables are not reported here. The F-statistic for the joint test of all the *YR*, *FIRM* variables is 5.484 (df=62, 602), which is significant at the 0.0001 level.

[b] Significant at the 0.01 level with one-tailed test.

[c] Significant at the 0.0001 level with one-tailed test.

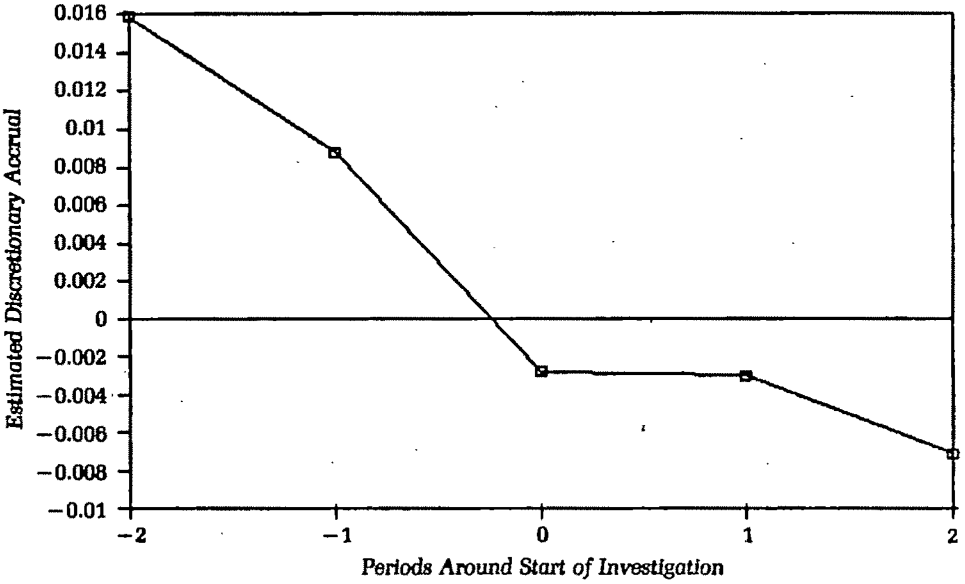
correlation between error terms. For equation (5), this statistic is 2.020 and is not significant from a positive or negative standpoint. This indicates that first-order autocorrelation is approximately zero.

The mean residual, or mean estimated discretionary accrual, is −0.0036 for the entire period of investigation and 0.0020 for the entire period of non-investigation. In figure 1, the residuals are plotted in five-year periods around the beginning and the end of the investigation. The residuals, which were 0.0159 two years before the investigation ($t_b=-2$), decreased to 0.0088 in the year before the investigation ($t_b=-1$), and declined again and became negative (−0.0028) in the year the investigation began ($t_b=0$). In the two years after the start of the investigation, the residuals were −0.0030 ($t_b=1$) and −0.0071 ($t_b=2$). The pattern shows a trend toward income-reducing discretionary accruals, with managers changing from net income increasing to net income decreasing accruals in the year the investigation began.

The residuals around the end of the investigation are expected to increase since presumably the firm's income was no longer being scrutinized by investigators. Two

Figure 1
Trends in Estimated Discretionary Accruals Around the Beginning and End of the Antitrust Investigation

Panel A. Trend for Years $t_b = -2$ *through* $t_b = +2$ *(Investigation began,* $t_b = 0$*):*



Panel B. Trend for Years $t_e = -2$ *through* $t_e = +2$ *(Investigation ended,* $t_e = 0$*):*

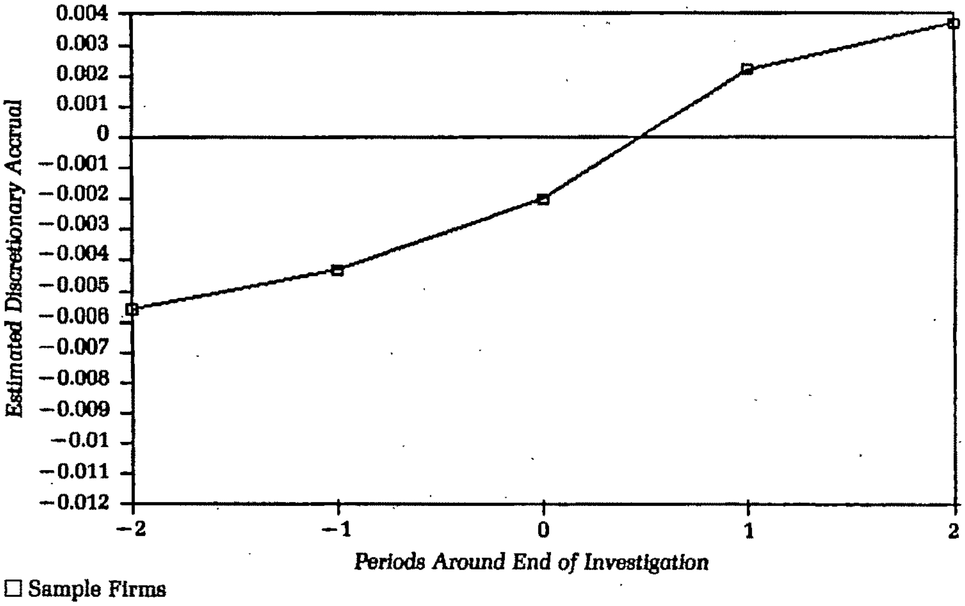


Table 2

Estimation Results for Equation (6) to Test the Effect of the Antitrust Investigation on Estimated Discretionary Accruals

$$A_{it}/TA_{it}=b_0[1/TA_{it}]+b_1[CHGSALES_{it}/TA_{it}]+b_2[FIXASSET_{it}/TA_{it}]$$
$$+b_3YR_{i1}+\ldots+b_{17}YR_{i13}+b_{18}FIRM_{1t}+\ldots+b_{64}FIRM_{47t}+b_{65}AT_{it}$$

Independent Variable[a]	*Expected Sign*	*Coefficient*	*t-value*
AT	−	−0.0108	−2.631[b]
CHGSALES/TA	+	0.0323	2.629[b]
FIXASSET/TA	−	−0.1046	−4.898[c]
1/TA		−0.2080	−0.425

Note: F-statistic = 18.969 (significant at 0.0001 with two-tailed test); adjusted $R^2=64$ percent; Durbin-Watson statistic = 2.043.

[a] Coding for *AT* variable: *AT* = 1 if investigation year; 0 otherwise. See table 1 for definitions of other regression variables and the estimation procedure. Coefficients and *t*-values for the dummy variables are not reported here. The *F*-statistic for the joint test of all the *YR*, *FIRM* variables is 5.620 (cf = 62, 601), which is significant at the 0.0001 level.

[b] Significant at the 0.005 level with one-tailed test.

[c] Significant at the 0.0001 level with one-tailed test.

years before the end of the investigation ($t_e=-2$), the residual was −0.0056, and one year before the end of the investigation ($t_e=-1$), the residual was −0.0043. The residual was −0.0020 in the year the investigation ended ($t_e=0$) and continued to increase in the next two periods to 0.0022 ($t_e=+1$) and 0.0037 ($t_e=+2$). The trend is consistent with accrual management but also implies that managers were able to anticipate the end of the investigation prior to its formal conclusion. Because the majority of investigations ended with either no lawsuit or a failed lawsuit, toward the end of the investigation, managers probably inferred that the government's position was weak and started taking more income-increasing accruals.[8]

Table 2 provides the regression results for equation (6), which tests statistically whether discretionary accruals were lower in the investigation period. The overall model is significant at the 0.0001 level and the adjusted R^2 increases to 64 percent. The *CHGSALES/TA* and *FIXASSET/TA* variables remain significant at the 0.005 level or better. The joint test for the dummy variables was significant at the 0.0001 level.

The added variable, *AT*, measures the incremental ability of the antitrust investigation to explain variation in the residuals from equation (5), the estimated discretionary accruals. Table 3 shows that *AT* is significant at the 0.005 level, with a one-tailed test, and its sign is negative. Because investigation years were coded as 1, the negative sign indicates that discretionary accruals in those years were lower than in other years, which supports the political-cost hypothesis.[9]

---

[8] Of the 48 firms, only 12 settled their cases, either voluntarily or involuntarily, in an unfavorable manner. All other cases were either dropped before an actual lawsuit was filed or the lawsuit was unsuccessful. There are two nonexclusive reasons for the large number of unsuccessful investigations. First, the enforcing agencies were unable to prove the existence of any "share monopolies" (see fn. 5). Second, the election of President Reagan caused a change in the enforcement process, leading to the dismissal of numerous cases.

[9] A number of firms had missing observations, but these were concentrated in 1969. Only five firms had missing data in years after 1969. To see whether the results were affected by missing data, equation (6) was run

The remainder of this section examines two alternative explanations for the results, transitory earnings and omitted variables.

Transitory Earnings Test

The reported findings could be attributed to changes in the transitory component of earnings rather than to earnings management (Beaver and Morse 1978). That is, the targeted firms might have been experiencing a period of high transitory earnings just prior to the investigation, which attracted the government's attention and led to the investigation. Because a firm with a high transitory component in one year would be expected to have lower earnings in the next year, the decrease in discretionary accruals after the start of an investigation might reflect a reversion to the mean.

With the assumption that the lengths of the transitory cycles are randomly distributed among sample firms, the transitory earnings explanation would be most valid when the investigation was relatively short. The longer the investigation, the more likely a period of high transitory earnings would recur during the investigation. These subsequent high transitory earnings would cancel out the preceding low transitory component and would cause the mean discretionary accrual for the period of investigation to regress toward some normal level. Hence, if the reported results of this study were caused by transitory earnings, the findings should be strongest for firms with short investigations. If the results were caused by earnings management, they should hold for all firms.

To discriminate between these two explanations, equation (6) was expanded to include a new variable, *ATL*. This variable was coded 1 for firm/year observations during the period of investigation if the firm's investigation was longer than 4.85 years, the average length for all firms. *ATL* examines whether the length of the investigation makes a difference in the *AT*-discretionary accrual relationship. If transitory earnings are driving the results, *ATL* will be positive and will cancel out the effect of *AT* for firms with long investigation periods.

This model was also significant at the 0.0001 level ($F=18.667$, $df=67$, 600), and *AT* is significant at the 0.01 level with a one-tailed test ($t=-2.285$). *ATL* is positively signed but is insignificant with a two-tailed test ($t=0.517$). Thus, responses were not different for investigated firms with long versus short investigation periods, which is inconsistent with the transitory earnings explanation.

Control Group Tests

The second alternative explanation is that the results were due to some omitted variable rather than the antitrust investigation. Accordingly, the tests were re-examined with a control group. The use of a control group is problematic because firms

---

for the 43 firms with complete data for years 1970–83. The results were essentially the same as those reported in table 3. To examine the robustness of the results, a series of sensitivity tests were performed. Because ten antitrust actions affected more than one firm, equation (6) was estimated ten additional times, excluding one group of firms at a time, to see whether the firms affected by one particular lawsuit were driving the results. Because the sample is skewed toward industries with multiple defendants, equation (6) was re-estimated by using industry portfolios instead of individual firms. The portfolios were formed by aggregating the data for all firms affected by the same antitrust investigation. This procedure gives equal weighting to each of the 24 investigations, as opposed to each firm, and examines whether the decrease in discretionary accruals was caused by a few investigations. The results of the sensitivity tests indicate the reported results are robust.

Table 3
Descriptive Statistics for Sample and Control Groups

Variable[a] and Group	Mean (s.d.)	*Median*	*Maximum*	*Minimum*	*t-value*[b]
A/TA					
Sample	−0.036 (0.050)	−0.037	0.173	−0.246	
Control	−0.031 (0.062)	−0.032	0.221	−0.238	−1.55
CHGSALES/TA					
Sample	0.111 (0.147)	0.107	0.810	−0.443	
Control	0.118 (0.231)	0.102	1.474	−1.957	−0.55
FIXASSET/TA					
Sample	0.796 (0.216)	0.792	1.628	0.216	
Control	0.739 (0.286)	0.711	1.690	0.163	4.03[c]
TA					
Sample	8270.9 (15781.1)	3071.3	149530.0	12.2	
Control	1286.7 (1785.6)	589.7	13094.0	4.3	11.56[c]

[a] Descriptive statistics reflect all firm/year observations for which data was available; see table 1 for definitions of variables. *TA* is in millions.

[b] The *t*-value is for the test that compares the sample mean with the control group mean. The significance levels reflect two-tailed tests. The positive *t*-values indicate the sample mean exceeded the control group mean.

[c] Significant at the 0.0001 level with two-tailed test.

being investigated for antitrust violations are not only the largest in their industry but also among the largest in the nation and world. Thus, it is not possible to match the sample firms on both industry and size. Nonetheless, because industry trends could be important in explaining patterns in discretionary accruals, an industry-matched control group was created by matching each sample firm with a randomly selected firm having the same four-digit SIC code (or three-digit when no full match was possible).

Table 3 provides descriptive statistics for the sample and control firms. As might be expected, the sample firms were both more capital-intensive and larger than the control firms. This significant difference for size provides support for use of the size proxy for at least one form of political cost, the cost associated with being the target of an antitrust investigation.

To provide insight on the year-to-year variation in the discretionary accruals of the control firms, equation (5) was estimated and the residuals were plotted in the five-year periods around $t_b = 0$ and $t_e = 0$, where t_b and t_e are the first and last years of investigation for the matched sample firm.[10] These residuals along with those for the sample firms (from figure 1) are shown in figure 2. The patterns of accruals are different. Sample firms' accruals exhibit a distinct trend. The overall pattern for the control firms is cyclical, with increases generally succeeded by decreases in the following period. This pattern indicates a lack of accrual management on the part of the control firms.

Figure 3 incorporates all years in the test period and shows the mean residuals for the entire period of pre-investigation, investigation, and post-investigation for both the sample and control groups. The residuals for the sample group show a large decrease during the investigation, while those for the control group remain almost constant.

The statistical significance of the different trends was evaluated in two ways. Equation (6) was first re-estimated with the control firms. This test examines whether the discretionary accruals of the control firms were statistically different when compared over the investigation and non-investigaton periods of their matched firms. The results in table 4 indicate that the model, control variables, and dummy variables are significant at the 0.005 level or better. AT, which was coded 1 for control firm i in year t if its matched sample firm was being investigated in year t, is incorrectly signed and, as expected, is not significant ($p=0.5826$) with a two-tailed test.

The second test used the discretionary accruals of the control firms as a benchmark measure of the "normal" level of discretionary accruals. The residual difference was computed for 558 firm/year observations with complete data, by taking the sample firm's residual minus its matched control firm's residual for the same period. A negative difference signifies that the sample firm had lower discretionary accruals than the control firm. The differences were examined to determine whether the difference was more negative in the investigation period. The mean difference is -0.0049 in the investigation period compared with a mean difference of 0.0025 in the non-investigation periods. With a t-test estimated by using unequal variances, this difference is significant at the 0.05 level on the basis of a one-tailed test ($t=2.060$). Thus, after controlling for industry trends, sample firms took significantly more income-reducing discretionary accruals while being investigated. The results of these two tests suggest that omitted, industry-related variables were not driving the results reported earlier in the study.[11]

---

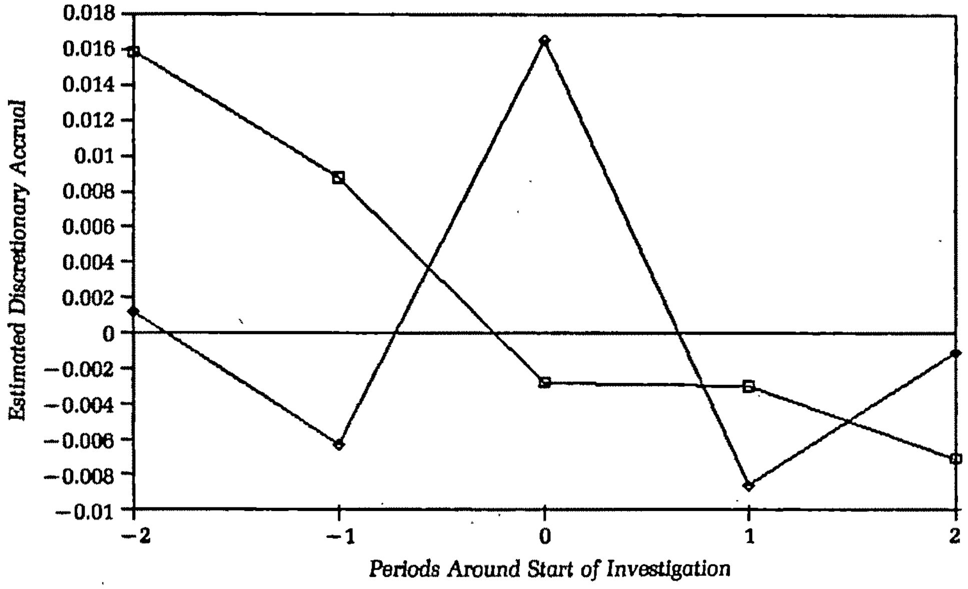
[10] The model estimated with the sample firms was able to explain considerably more variation in deflated total accruals than the model estimated with the control firms (adjusted R^2 equal to 63.7 and 38.9, respectively). There are a number of possible reasons why this might occur: (1) The control firms may take relatively more discretionary accruals than the sample firms. This could occur if control firms use more liberal accounting methods that increase the managers' flexibility. (2) The control firms have relatively less depreciation. Table 4 shows that the control firms were less capital-intensive than the investigated sample. (3) The model is misspecified for control firms. There may be additional variables affecting nondiscretionary accruals that are particularly important to the control firms.

[11] Because the DOJ and FTC focus on return on assets (ROA), ROAs were also compared for the sample and control firms. Although the sample firms had significantly higher ($p<0.10$) reported ROAs in the pre-investigation period, the difference was not significant once an investigation had begun. Thus, while sample firms appeared more profitable prior to investigation, their profitability was comparable to their industry during the investigation. Rather, if the effects of the discretionary accruals are ignored, the sample firms' ROAs would have also been higher in the investigation period, and this difference would be marginally significant at the 0.10 level. This implies that the comparability of reported ROAs in the investigation period was at least partly due to the negative discretionary accruals taken by the sample firms.

Figure 2

Trends in Estimated Discretionary Accruals Around the Beginning and End of the Antitrust Investigation for Sample and Control Firms

Panel A. Trend for Years $t_b = -2$ *through* $t_b = +2$ *(Investigation began,* $t_b = 0$*):*



Panel B. Trend for Years $t_e = -2$ *through* $t_e = +2$ *(Investigation ended,* $t_e = 0$*):*

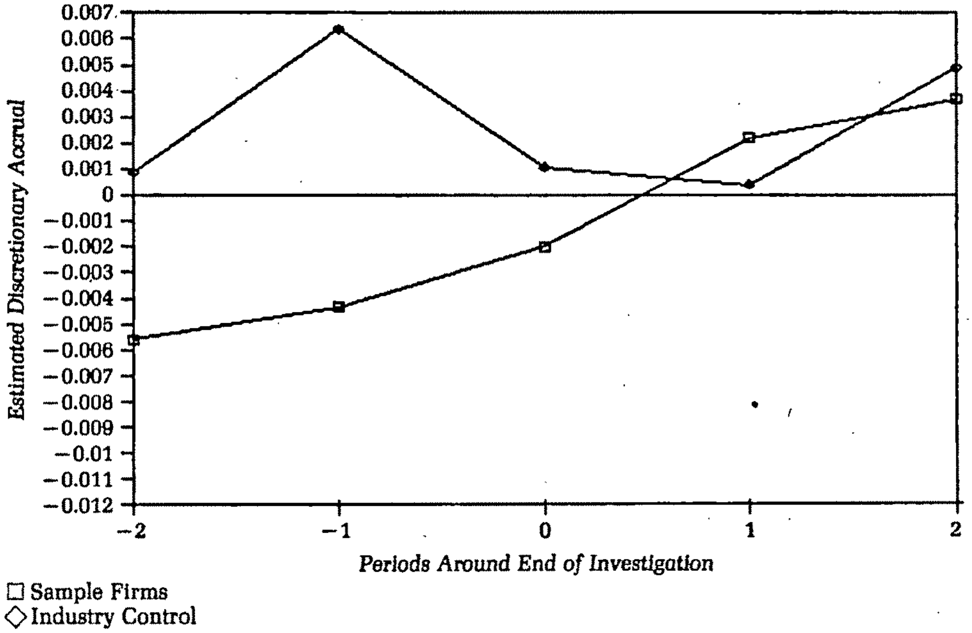


Figure 3

Trends in Estimated Discretionary Accruals From Pre- to Post-Investigation for Control and Sample Firms

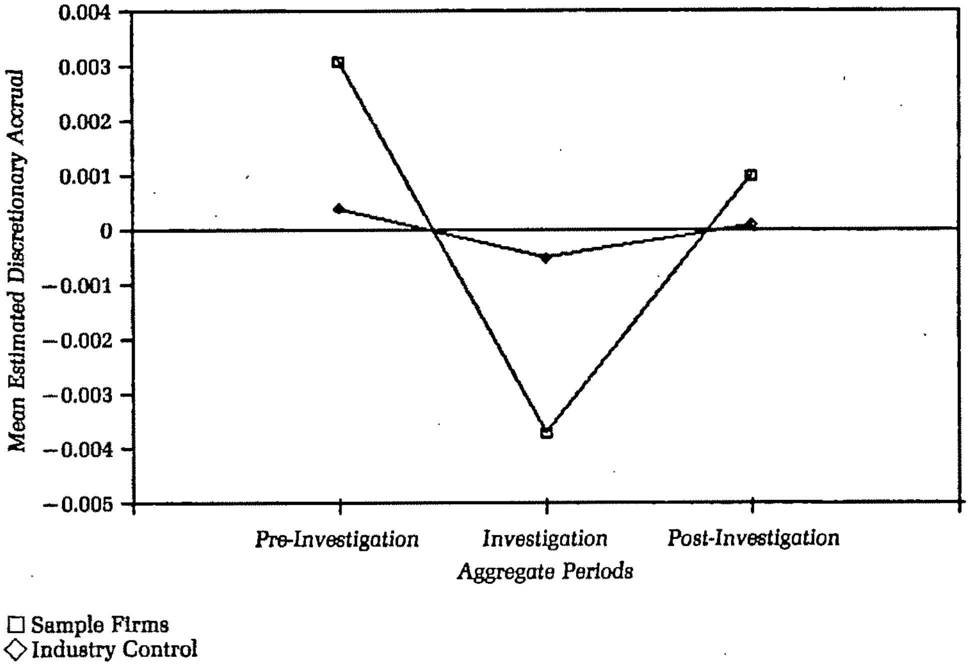


Table 4
Estimation Results for Equation (6) With Control Sample

$$A_{it}/TA_{it}=b_0[1/TA_{it}]+b_1[CHGSALES_{it}/TA_{it}]+b_2[FIXASSET_{it}/TA_{it}]$$
$$+b_3YR_{i1}+\ldots+b_{17}YR_{i13}+b_{18}FIRM_{1t}+\ldots+b_{64}FIRM_{47t}+b_{65}AT_{it}$$

Independent Variable[a]	Expected Sign	Coefficient	t-value
AT	−	0.0034	0.550
CHGSALES/TA	+	0.0354	2.890[b]
FIXASSET/TA	−	−0.0537	−2.540[b]
1/TA		−0.6227	−1.617

Note: F-statistic=7.434 (significant at 0.0001 with two-tailed test); adjusted R^2=38.94 percent; Durbin-Watson statistic=2.170.

[a] See tables 1 and 2 for definitions of other regression variables. The model was estimated using least squares regression for 666 observations related to 48 firms in the industry-matched control group for the years 1969–83. Coefficients and t-values for the dummy variables are not reported here. The F-statistic for the joint test of all the YR, FIRM variables is 2.451 (df=62, 600), which is significant at the 0.001 level.

[b] Significant at 0.005 level with one-tailed test.

V. Conclusion

This study examines the effect of monopoly-related antitrust investigations on firms' reported earnings. The evidence, which indicates that managers adjust their discretionary accruals in response to changes in potential political costs, supports the political-cost hypothesis.

The results, however, are limited in two ways. First, the study focuses on the reaction of firms to one type of political action. Thus, the results may not be generalizable to other political actions, especially if the link between accounting income and the political decision is not well defined. Second, the study examines how firms react to investigation once it occurs but does not specifically address what triggers investigation or how firms behave in an attempt to preclude such investigations.

These limitations suggest some areas for future research. For example, the longitudinal approach could be extended to explore responses to a wide range of political activities (e.g., congressional legislative actions, other regulatory actions). Also, while developing proxies for the firm's exposure to future political actions is difficult, such research is clearly warranted.

Appendix
Sample Firms Used in Study

SIC	*Firm*	*Monopolized Industry/Product*	*Investigation Began/Ended*		*Matched Firm*[1]
1000	Cleveland-Cliffs	Iron ore	1977	1980	INCO
2000	Quaker Oats	Breakfast cereal	1970	1978	Unilever
2020	Borden	Lemon juice	1974	1978	Kraft
2040	General Mills	Breakfast cereal	1970	1982	International Multifoods
2040	Kellogg	Breakfast cereal	1970	1982	Ralston Purina
2800	Dow Chemical	Magnesium	1977	1979	American Cynamid
2820	DuPont	Titanium dioxide	1978	1980	Grace (W.R.)
2821	Union Carbide	Dry cell batteries	1977	1981	Goodrich (B.F.)
2821	Hercules	Nitrocellulose	1979	1981	Rohm & Haas
2834	Bristol-Myers	Ampicillin	1970	1981	Abbott Labs
2834	Lilly (Eli)	Cephalosprin	1975	1977	Upjohn
2840	Colgate-Palmolive	Detergent	1975	1980	Ecolab
2840	Proctor & Gamble	Detergent	1975	1980	Stanhome
2911	AMOCO	Oil refining	1973	1981	Crown Central
2911	Atlantic Richfield	Oil refining	1973	1981	Holly
2911	Chevron	Oil refining	1973	1981	Imperial Oil
2911	Exxon	Oil refining	1973	1981	Kerr-McGee
2911	Gulf	Oil refining	1973	1981	Pauley Petroleum
2911	Mobil	Oil refining	1973	1981	Phillips
2911	Shell	Oil refining	1973	1981	Quaker State
2911	Texaco	Oil refining	1973	1981	Sun Oil
3011	Goodyear	Tire	1973	1976	Alliance Tire
3011	Firestone	Tire	1973	1976	Armtek
3069	Hanna Mining	Iron ore	1977	1980	Vulcan
3312	Armco	Flat-rolled steel	1977	1980	Inland Steel
3312	Bethlehem Steel	Flat-rolled steel	1977	1980	Lukens
3312	U.S. Steel	Flat-rolled steel	1977	1980	Texas Industries
3312	Wheeling-Pittsburg	Flat-rolled steel	1977	1980	Quanex
3330	Alcan Aluminum	Aluminum	1975	1980	AMAX

Appendix—*Continued*

SIC	Firm	Monopolized Industry/Product	Investigation Began	Investigation Ended	Matched Firm[1]
3330	ALCOA	Aluminum	1975	1980	Asarco
3330	Kaiser Aluminum	Aluminum	1975	1980	Brush Wellman
3330	Phelps Dodge	Aluminum	1975	1980	Cominco
3330	Reynolds Metals	Aluminum	1975	1980	O'Okiep Copper
3540	Black & Decker	Power tools	1979	1980	Acme Precision
3570	Pitney Bowes	Postage meters	1976	1977	Rockaway
3600	General Electric	Generators	1976	1977	Motorola
3600	RCA	Network TV	1972	1976	Texas Instruments
3664	Westinghouse	Generators	1972	1976	Litton Industries
3711	Chrysler	Automobile	1976	1981	Federal Signal
3711	Ford	Automobile	1976	1981	Honda
3711	General Motors	Automobile	1976	1981	Paccar
3842	Everest & Jennings	Wheelchairs	1977	1979	Becton Dickinson
3861	Xerox	Copiers	1972	1975	Polaroid
3949	Bally	Slot machines	1978	1983	Caesars World
4811	ATT	Telecommunications	1973	1982	S. New England Telecomm
4833	ABC	Network TV	1972	1976	Lin Broadcasting
4833	CBS	Network TV	1972	1976	Communication Satellite
4891	Viacom	TV programs	1972	1973	Acton

[1] Matched firms are used in subsequent tests.

References

Areeda, P. 1981. *Antitrust Analysis.* 3d ed. Boston: Little, Brown.

Ball, R., and G. Foster. 1982. Corporate financial reporting: A methodological review of empirical research. *Journal of Accounting Research* 20 (Supplement): 161–234.

Beaver, W., and D. Morse. 1978. What determines price-earnings ratio? *Financial Analysts Journal* 34 (July–August): 65–76.

Bowen, R., E. Noreen, and J. Lacey. 1981. Determinants of the corporate decision to capitalize interest. *Journal of Accounting & Economics* 3 (August): 151–79.

DeAngelo, L. 1986. Accounting numbers as market valuation substitutes: A study of management buyouts of public stockholders. *The Accounting Review* 61 (July): 400–20.

Elzinga, K. 1989. Unmasking monopoly: Four types of economic evidence. In *Economics and Antitrust Policy,* edited by J. Meehan and R. Larner. New York: Quorum Books.

Fisher, F., J. McGowan, and J. Greenwood. 1983. *Folded, Spindled, and Mutilated: Economic Analysis and U.S. v. IBM.* Cambridge, MA: MIT Press.

Healy, P. 1985. The effect of bonus schemes on accounting decisions. *Journal of Accounting & Economics* 7 (April): 85–107.

Howard, M. 1983. *Antitrust and Trade Regulation: Selected Issues and Case Studies.* Englewood Cliffs, NJ: Prentice-Hall.

Jones, J. 1990. Earnings management during import relief investigations. Working paper, University of Chicago, IL.

Judge, G., W. Griffiths, R. Hill, H. Lutkepohl, and T. Lee. 1985. *The Theory and Practice of Econometrics.* 2nd ed. New York: John Wiley & Sons.

Liberty, S., and J. Zimmerman. 1986. Labor union contract negotiations and accounting choices. *The Accounting Review* 61 (October): 692–712.

McNichols, M., and P. Wilson. 1988. Evidence of earnings management from the provision for bad debts. *Journal of Accounting Research* 26 (Supplement): 1–31.

Meehan, J., and R. Larner. 1989. The structural school, its critics, and its progeny: An assessment. In *Economics and Antitrust Policy,* edited by J. Meehan and R. Larner. New York: Quorum Books.

Moyer, S. 1990. Capital adequacy ratio regulations and accounting choices in commercial banks. *Journal of Accounting & Economics* 13 (July): 123–54.

Peltzman, S. 1976. Toward a more general theory of regulation. *Journal of Law and Economics* 19 (August): 211–40.

Pindyck, R., and D. Rubinfeld. 1981. *Econometric Models and Forecasts.* 2nd ed. New York: McGraw-Hill.

Schipper, K. 1989. Earnings management. *Accounting Horizons* 3 (December): 91–102.

Scholes, M., P. Wilson, and M. Wolfson. 1989. Tax planning, regulatory capital planning, and financial reporting strategy for commercial banks. Working paper, Stanford University, CA.

Stigler, G. 1971. The theory of economic regulation. *Bell Journal of Economics and Management Science* 2 (Spring): 3–21.

Sutton, T. 1988. The proposed introduction of current cost accounting in the U.K.: Determinants of corporate preference. *Journal of Accounting & Economics* 10 (April): 127–49.

Watts, R. 1977. Corporate financial statements, a product of market and political processes. *Australian Journal of Management* 2 (April): 53–75.

———, and J. Zimmerman. 1978. Toward a positive theory of determination of accounting standards. *The Accounting Review* 53 (January): 112–34.

———, and ———. 1986. *Positive Accounting Theory.* Englewood Cliffs, NJ: Prentice-Hall.

Wong, J. 1988. Political costs and an intraperiod accounting choice for export tax credits. *Journal of Accounting & Economics* 10 (January): 37–51.

Zimmer, I. 1986. Accounting for interest by real estate developers. *Journal of Accounting & Economics* 8 (March): 37–51.

A FORUM ON

LABORATORY MARKETS AND AUDITING RESEARCH

1. Nicholas Dopuch and Ronald R. King
   Negligence Versus Strict Liability Regimes in Auditing: An Experimental Investigation

2. David E. Wallin
   Legal Recourse and the Demand for Auditing

3. *The Arizona State University Research Workshop*
   James R. Boatsman, Lawrence P. Grasso, Michael B. Ormiston, and J. Hal Reneau
   A Perspective on the Use of Laboratory Market Experimentation in Auditing Research

4. *The University of Iowa Accounting Workshop*
   Douglas V. DeJong, and Robert Forsythe
   A Perspective on the Use of Laboratory Market Experimentation in Auditing Research

---

Editor's Note: A "Forum" is a collection of articles around a common theme. Accepted refereed articles are followed by an invited "perspective" authored by experts on that theme. In addition to allowing for synergy and providing alternative explanations for the results, an objective of the "perspective" is to provide guidance for future research on issues related to that common theme.

THE ACCOUNTING REVIEW
Vol. 67, No. 1
January 1992
pp. 97–120

Negligence Versus Strict Liability Regimes in Auditing: An Experimental Investigation

Nicholas Dopuch
Ronald R. King
Washington University

SYNOPSIS AND INTRODUCTION: In this study we assess how different regimes of auditor liability affect the demand for and supply of auditing services. The assessment was made with 15 experimental markets, each of which involved two sellers of assets, two auditors (verifiers), and four buyers. The experimental markets paradigm allowed us to compare the negligence liability regime (six markets) that auditors currently face with two alternatives not currently in existence—a strict liability regime (six markets) and a no-liability regime (three markets). We focused on the extent to which the experimental results conformed to our predictions of (1) sellers' frequency of hiring verifiers and of selecting a costly investment that improved aggregate welfare, (2) verifiers' service fees and their frequency of testing the truthfulness of the sellers' disclosures, and (3) buyers' reliance on the sellers' disclosures and verifiers' reports when pricing the sellers' assets. The predictions varied across the liability regimes primarily because differences in the degree of the verifiers' liability changed their economic incentives to test the truthfulness of the sellers' disclosures.

The results show that the no-liability and negligence markets operated in a manner consistent with the predictions, whereas the strict markets deviated from the predictions on several dimensions. Specifically, verifiers in the strict liability markets were hired less often than predicted because they submitted higher offers for their services than sellers were willing to pay. This in turn led to fewer than predicted costly investments by the sellers.

The assistance and comments of Jon Davis, Doug DeJong, Thomas Fields, Carol Frost, Mahendra Gupta, Mort Pincus, Ramji Balakrishnan, Chuck Schnitzlein, David Wallin, and seminar participants at the Universities of Florida, Iowa, Ohio State, and North Carolina and the 1990 AAA meetings are acknowledged. Financial support was provided by the KPMG Peat Marwick Foundation Research Opportunities in Auditing and Research Fellowship Programs.

Submitted March 1990.
Accepted August 1991.

Although our general conclusion recognizes that a legal system is an integral part of the auditing institutional infrastructure, we found no evidence of any systematic benefits from imposing a strict liability rule on the verification service. In fact, the results suggest that the negligence liability markets operated at a level of economic efficiency as high or higher than those in the other two regimes. This suggests that the current tendencies of courts and the auditing profession to expand the scope of auditors' liabilities may not achieve the net benefits expected from such expansions.

Key Words: *Auditing, Strict liability, Negligence liability.*

Data Availability: *The experimental data generated in this study may be obtained from the authors upon request.*

IN the next section, we provide some background discussion of liability issues in auditing, along with a general description of our experimental markets. Details of our experimental design appear in section II. Hypotheses about the strategies of sellers, verifiers, and buyers are derived in section III, and results from our tests of these hypotheses are discussed in section IV. Conclusions are provided in section V.

I. Litigation Concerns in Auditing

Basic Liability Issues in Auditing

Congressional oversight committees appear to view recent increases in auditors' liability as evidence that the industry has not been providing an adequate service.[1] In contrast, many in the accounting profession argue that increases in liability reflect symptoms of the present legal system that could jeopardize the viability of the profession.[2] In an effort to exert some control over increased legal exposure, the profession has promulgated additional auditing standards, worked to narrow the "expectations gap," and recommended certain statutory changes.[3] Some of the new standards, however, may actually increase liability exposure by extending auditors' responsibilities to services not previously considered a legal basis for assessing auditors' liability (e.g., extending auditors' responsibility for detecting managerial fraud).

For purposes of discussion, we present two main dimensions of auditors' liability. The first dimension relates to four categories of parties who may sue the auditor. These are: (1) client parties, (2) primary beneficiaries (parties for whose benefit the audit was performed), (3) foreseen parties (parties the auditors should have reasonably foreseen would use the audit report), and (4) foreseeable parties (parties the auditor could have expected to use the report).

The second dimension of auditors' liability relates to the degree of care for which the auditor could be held liable in providing auditing services. The four levels are fraud, gross negligence (constructive fraud), negligence, and finally responsibility for conducting audits without errors. The degree of care required of an auditor depends upon both the plaintiffs' legal standing and their respective legal bases (contractual, tort, or security law).

---

[1] See, for example, Minow (1984), Mednick (1987), and Cowan (1990).

[2] Recently, Laventhol and Horwath filed for bankruptcy, citing their high liability costs as one cause. See Pae (1990).

[3] For example, the AICPA supports the "Lawsuit Reform Act of 1989" (S.1100) which proposes to significantly modify the tort system.

Historically, auditors were held liable to foreseen or foreseeable parties for fraud and gross negligence, but not for negligence alone.[4] The due-care standard serves as the measure of the degree of care separating the negligence and the errorless categories. The cutoff has been defined either at the point where the expected costs of audit and litigation are minimized, or at the level specified as meeting auditing standards.[5] Of course, it is often difficult for auditors to know whether a particular audit will meet the due-care standard as defined by the courts when attempting to design an optimal audit program. Expanding auditors' responsibilities to include requirements of explicit reports on internal control, going concern status, and managerial fraud make that determination even more difficult since there are presently few precedents that can provide guides to the possible interpretations of these standards by the courts.

The classifications given above may be used as a point of reference in differentiating the three liability regimes we examine in this study. The no-liability regime represents the case in which the verifier is not liable regardless of the level of services provided. This regime will be used as a benchmark for assessing the effects of the other two liability regimes. In the strict liability regime, verifiers are held liable, given that a loss occurred for some party, regardless of the service level. In such a regime, the degree of care does not eliminate qualified parties' standing to sue. In the negligence liability regime, verifiers are not liable if they provide the due-care level of service. This regime most closely resembles the one under which auditors currently operate.

Although auditors presently do not operate under a strict liability regime, we believe the distinction between a negligence and a strict liability regime has become blurred by the expansion of the number of parties who have a standing to sue and by increases in the scope of auditors' liability into areas where the due-care standard may be difficult to define. These changes have paralleled those observed with product liability, and they have moved some courts to suggest that auditors even now have an implied warranty of results similar to a strict liability standard (see Cave 1985).[6]

Proponents of a strict liability regime argue that it provides (potential) defendants with incentives to increase the quality of their products and services, thereby lowering the probability of injuries or losses.[7] However, imposing strict liability on producers may reduce the incentives for users and consumers to exercise reasonable care in how they use products and services.[8] Moreover, it may be difficult for producers of some goods and services to estimate the potential costs of their output failures since these costs will be influenced by factors largely under the control of plaintiff users.[9] For example, auditors may be held liable for failure to predict business failures even when the managers of audited firms subsequently undertake highly risky projects.

---

[4] The extent to which foreseen and foreseeable third parties have standing presently depends on the state in which the suit is brought. Many articles have appeared recently in legal journals discussing the pros and cons of holding negligent auditors liable to foreseen and foreseeable parties. See Kothari et al. (1988) and Gormley (1985) for a comprehensive review. The evolution of the definition of the privity condition in auditing, beginning with the well-known Ultramares Corp. v. Touche, Niven & Co. case, is summarized in Guy et al. (1990).

[5] DeJong and Smith (1984) find evidence that courts set due-care standards at the minimum of the combined audit and litigation costs. However, Adler (1989) suggests that juries have a difficult time in resolving complex cases wherever the standard is set. In most court cases auditors have been found to meet the due-care standard if they meet or exceed the professional standards, but not always. For example, see United States v. Simon, 425 F2d 796 (2ndCir. 1969), cert denied, 397 U.S. 1006 (1970), discussed in Kinsella (1984).

[6] For example, in Rosenblum v. Adler the court argued that auditors are the most appropriate bearers of risk of loss because they can pass the costs to all clients and can purchase insurance (Rosenblum, Inc. v. Adler, 93 N.J. 324, 461 A.2d 138, 1983).

[7] See Chilton (1990) for a discussion of product liability reform.

Overview of Experimental Design

Since the focus of this study is on comparisons of existing and contemplated institutional arrangements where no naturally occurring data exists, we use the method of experimental economics.[10] Potential effects are compared across 15 experimental markets: three in the no-liability regime and six each in the negligence and strict liability regimes. A simple game-theoretic model is used to form predictions across the different regimes.

Our basic setting for all 15 markets includes a seller of a dividend-paying asset transacting with a group of potential buyers. Prior to receiving an asset, the seller decides whether to make a costly investment that increases the probability of having a high-quality asset available for sale. The seller's expected return from the costly investment depends on the buyers' beliefs about which asset quality the seller actually has for sale. The seller makes a disclosure to the buyers about the asset's quality, but these disclosures may not be credible because sellers could lie. Hence, buyers are predicted to ignore disclosures by assuming that sellers will misrepresent low-quality assets as high ones. This, in turn, will reduce sellers' expected returns from the costly investment and, thus, their willingness to make that investment.

Dopuch and King (1991) observed that a computerized verification of information provided sufficient credibility to sellers' disclosures for the buyers to condition their bids on those disclosures. This, in turn, allowed the sellers to earn a positive return from the costly investment. The net effect was a significant increase in the frequency of sellers' costly investments (and in economic welfare) in markets with computerized verifiers.

However, these results were not obtained when human verifiers subject to the moral-hazard problem of failing to perform their service were hired to verify the sellers' disclosures. Their verification reports lacked credibility, and this resulted in a decrease in both the hiring of verifiers and costly investments.[11] However, the verifiers in that study were not subject to any legal penalties for nonperformance of their service.

Using a setting similar to that of our previous study, we investigate here the extent to which either a negligence or a strict liability regime would alleviate verifiers' moral-hazard problem. In either setting, verifiers hired by sellers could be sued for damages under certain conditions by buyers who have purchased an asset that had a zero payoff. In the negligence regime, verifiers were absolved of any liability for damages if they had performed their service because the performance of the service was deemed equivalent to satisfying the due-care standard. In contrast, verifiers hired in the strict liability regime were held liable for damages under the same conditions, even if the verification service had been performed. In both liability regimes, the damages paid by the verifiers consisted of the difference between the asset's actual price and its expected value as a low-quality asset.[12]

---

[8] Most states have comparative negligence rules rather than contributory negligence rules. See Prosser and Keeton (1984), Shavell (1980), and Posner (1975a, 1975b).

[9] This has often led to the use of "caps" or upper limits on damages suffered by consumers who are compensated under a strict liability regime.

[10] Two previous experimental studies of alternative liability regimes within auditing contexts are DeJong et al. (1985) and Wallin (1992). However, unlike the present study, neither of these involved strategic verifiers subject to moral-hazard problems. Illustrations of the use of experimental markets to study other auditing issues are Davis (1989) and Kachelmeier (1990). See Smith et al. (1987) for a discussion of the usefulness of experimental markets in auditing research. See Melumad and Thoman (1990) for a theoretical analysis of auditors and courts.

[11] The incidence of nonperformance was lower when verifiers were hired by buyers, but it still occurred.

Sellers were not liable for false disclosures in any of the regimes. One reason we did not include a liability rule for sellers is that Wallin (1992) uses a similar setting to investigate how a legal liability rule for sellers would affect their investment and disclosure decisions. Another reason is that auditors have had to pay all of the damages in a number of security cases even though managers were also guilty of negligence and/or fraud (the "deep pockets" syndrome).[13]

II. Experimental Design

Subjects and Procedures

A total of 120 subjects (undergraduate business students) participated in the 15 markets (eight subjects in each market). Subjects volunteered for two different sessions scheduled several days apart. At the first session, subjects received written instructions that were read aloud by an assistant.[14] After clarifying questions were answered, participants were seated at computer terminals and given instructions that familiarized them with the operation of the computer and allowed them to practice transacting via computer by making decisions in all three agent roles (sellers, buyers, and verifiers). This first session of extensive instructions was scheduled because of the relative complexity of our markets. Previous experimental research has demonstrated that the volatility of subjects' behavior decreases with experience in the particular market settings and with exposure to decisions faced by other players (Forsythe and Lundholm 1990). Each subject's pay for the first session was based on the score received on a quiz given at the end of the session (cash payments averaged $8.25 out of a maximum of $10).

When the participants returned for the second session, they again received instructions for a particular market treatment and then went through the computer instructions for all three roles a second time. Subjects were then randomly assigned to roles, and the market periods were conducted. Upon completion of all periods (the steps making up a period will be described later), subjects completed a questionnaire, were paid (privately) their earnings in cash, and dismissed. The participants never learned which individuals played the other roles.[15] Subjects indicated in a questionnaire that they had fully understood the instructions for the second session. The questionnaire also revealed that subjects felt adequately compensated (average earnings were $35 for this session, which took less than two hours) and that they understood how their decisions affected their cash payments.

---

[12] The party that lost the suit also had to pay an enforcement (court) cost in addition to the damages.

[13] This has led to proposals by the profession that states adopt a several liability rule in cases involving auditors (see "Washington Update" 1989). Currently, most states still follow the joint and several rule in civil liability cases.

[14] Instructions are available from the authors upon request.

[15] The computer laboratory had physical barriers that prevented participants from making visual contact with each other. Verbal contact was prohibited and monitored by the administrator.

Steps for Each Period

Each of our markets involved two sellers, two verifiers, and four buyers who transacted for 20 periods.[16] All market communications and exchanges were handled by a system of networked personal computers. Table 1 summarizes our notation and the parameter values used. The following discussion of the steps comprising each market period will be illustrated with a hypothetical example of decisions that agents could make.

Step 1. Each verifier (denoted as V1 and V2) submits a verification offer to each seller. Each offer is the amount a seller must pay to the verifier if that verifier is hired for the verification service. In the present hypothetical example, V1 submits an offer of \$0.65 to each seller (S1, S2) and V2 submits offers of \$0.64 and \$0.60 to S1 and S2, respectively.

Step 2. Each seller decides which verifier (if any) to hire. In the present example, S1 chooses to hire V1 for verification at the offer of \$0.65, whereas S2 chooses to hire V2 for that service at \$0.60. Sellers pay the amount of the accepted offers whether or not the verifiers perform their tests.

Step 3. The sellers make their investment decisions. This was operationalized by having the sellers choose between a cost-free investment (I_N) or a costly investment (I_Y), the latter requiring the sellers to pay \$0.75 to the experimenter.[17] If I_Y is selected, the seller has a 0.80 probability of receiving a high-quality (q_H) asset and a 0.20 probability of receiving a low-quality (q_L) asset. If I_N is selected, the seller has a 0.20 probability of receiving a q_H asset. A q_H asset has a 0.90 probability of a payoff value of \$4 to the purchasing buyer, and a 0.10 probability of a zero payoff. A q_L asset has a 0.10 probability of a \$4 payoff and a 0.90 probability of a zero payoff.[18] In the example, S1 chooses I_N and S2 chooses I_Y. Thus, S1 has a 0.20 probability of receiving a q_H asset, whereas S2 has a 0.80 probability of receiving a q_H asset. Assume in this example that S1 receives a q_L asset and S2 receives a q_H asset.

Step 4. Each seller learns the quality of his or her asset and selects a disclosure (which is not revealed to the verifier until after the testing decision is made). Sellers were required to disclose q_H assets as high; however, q_L assets could be either falsely disclosed as high or truthfully as low.[19] Prior to receiving a seller's disclosure, the hired verifier decides whether to pay \$0.50 for a test (select T_Y) which would inform the veri-

---

[16] The market setting and procedures were selected in order to have the simplest environment that captured some basic frictions of an auditing market. The competition between the two verifiers can be characterized as Bertrand competition, which only requires two competitors. Four buyers were used because previous experimental economics studies show that four bidders generate competitive outcomes (Dopuch and King 1991; Wallin 1992).

[17] The cost of investing was set at \$0.75 to be high enough to be a salient amount, but low enough that the expected economy-wide resources were greater when I_Y was selected than when I_N was selected (after considering the \$0.75 cost of selecting the costly investment).

[18] The compound setting allowing both asset qualities to have the potential to pay \$4 and allowing both investment choices to produce high assets was incorporated so that buyers could not invert to determine the investment selection of the seller.

[19] The requirement that sellers had to disclose high assets as high simplified the derivation of our hypotheses and was imposed because we found in previous markets that sellers disclosed q_H assets as low only five times out of 750 opportunities (Dopuch and King 1991). Thus, this requirement should not have been restrictive.

Table 1
Summary of Notation and Parameter Values

Asset Qualities		
High-quality (q_H)	=	90 percent chance of a $4 payoff to buyers and a 10 percent chance of a zero payoff.
Low-quality (q_L)	=	10 percent chance of a $4 payoff and a 90 percent chance of a zero payoff.
Sellers' Decisions		
Hiring of verifier (V_Y or V_N)	=	if hired, verifier is paid the offer amount by the seller and produces a verification report.
Costly investment (I_Y)	=	produces a distribution of 80 percent q_H assets and 20 percent q_L assets; cost of I_Y is $0.75.
Cost-free investment (I_N)	=	produces a distribution of 20 percent q_H assets and 80 percent q_L assets.
Verifiers' Testing		
Costly test (T_Y)	=	generates a truthful report about asset quality and agrees or disagrees with seller's disclosure; cost of T_Y is $0.50.
Cost-free test (T_N)	=	generates report of agreement with seller's disclosure.
Legal Regimes		
No-liability regime	=	no liability to any party.
Negligence liability regime	=	buyers can sue, if: (1) seller discloses high, (2) a verifier is hired, (3) zero payoff obtained, and (4) the asset price is greater than $0.55.
Strict liability regime	=	Suits are automatically implemented given the same four conditions listed under the negligence regime.
Legal costs	=	$1 for each suit.
Legal cost allocation	=	Legal costs are paid by the party losing the suit.

fier of the actual quality of the asset and, hence, the truthfulness of the seller's subsequent disclosure.[20] The disclosure and the verification report are then sent to the buyers. The report could reveal the actual asset quality to the buyers, but only if the verifier selected T_Y. The report could not disagree with the seller's disclosure unless T_Y was selected, nor could the verifier send a false report after selecting T_Y (e.g., by reporting agree when they knew the seller misrepresented the asset quality).[21] Only the verifier knows at this point if T_Y or T_N is selected.

[20] The $0.50 cost was selected to make the amount salient yet have economic surplus greater when costly investment and testing (I_Y and T_Y) were selected than when not.

[21] In effect, verifiers could issue false reports only if they did not perform their service, equivalent to deliberately failing to meet generally accepted auditing standards.

Because S1 has a q_L asset, the disclosure can be either high or low. If S1 falsely discloses high and V1 selects T_Y, the report is disagree, and buyers know the seller has a q_L asset; however, if S1 discloses high and the verifier selects T_N, the report must be agree and buyers will not be certain of asset quality. If S1 discloses low, the buyers again know the asset is q_L since sellers cannot disclose q_H assets as low. Hence, whenever the seller discloses low, the verification report will always agree. Similarly, whenever the verifier selects T_N, the verification report sent to the buyers is always agree.

In the example, S1 falsely discloses high and V1 chooses T_N, which generates the disclosure/report message combination of H/A. Since S2 has a q_H asset, the disclosure must be high, which will be confirmed by an agree report regardless of V2's test decision.

Step 5. The buyers—who now know (1) which verifier (if any) was hired for the verification service, (2) the sellers' disclosures, and (3) the verification reports—submit bids for the assets. Note that buyers in the example do not know whether the verifiers chose T_Y or T_N, or whether the sellers chose I_Y or I_N. Similarly, they do not know the truthfulness of the seller's disclosure. The buyer submitting the highest bid is awarded the asset at the second highest price. This second-price, sealed-bid institution was selected because it is theoretically fully demand-revealing.[22] Assume that the four buyers bid $2.20, $2.60, $3, and $1 for S1's asset and $2, $2.40, $1.90, and $2.40 for S2's asset. S1's asset is transferred to B3 who bid $3 at the second highest price of $2.60, and all bids are revealed to all agents. Since both B2 and B4 submitted the highest bid for S2's asset ($2.40), the computer randomly selects between them (assume B2 is selected and pays $2.40 for the asset).

Step 6. The asset's payoff ($4 or zero) is determined and communicated to all agents. Assume, for the purposes of the example, that B3 gets zero dollars and B2 gets $4. The period terminates at this point in the no-liability markets, but not necessarily in the negligence and strict liability markets.

Under the negligence regime, a purchasing buyer could decide to sue the verifier under the following conditions: (1) the seller disclosed high, (2) a verifier was hired (and presumably reported agree), (3) the asset payoff was zero, and (4) the purchasing buyer paid more than $0.55 or the asset.[23] If that buyer chooses to sue, the verifier's test decision (T_Y or T_N) is revealed to everyone. A verifier who selected T_Y would not have to pay any damages, and the buyer would have to pay $1 for enforcement costs.[24] But a verifier who selected T_N would have to pay damages to the buyer—equal to the asset's purchase price minus \$0.40—plus the \$1 enforcement costs.[25] In our example, B2 could not sue because the asset paid \$4. However, B3 could sue and is here assumed to do so. V1 then has to pay \$1.90 (\$2.40 less \$0.50) to B3 plus \$1 of enforcement costs to the experimenter. All market participants are informed of the outcome of the suit.

---

[22] The expected utility-maximizing bid of the second-price sealed-bid institution is equal to the value determined by the bidder and does not depend on the bidder's expectation of other players' bids. See Cox et al. (1982) for a discussion of this institution.

[23] The $0.55 minimum was imposed to reduce the frequency of automatic suits in the strict liability regime. This minimum should not have had any systematic effect in the negligence regime. This can be seen by noting that there are three possible disclosure/report combinations: high/agree, high/disagree, and low/agree. In the case of high/agree, buyers are predicted to bid approximately $3.50 for the asset. In the other two cases the expected bids are approximately $0.40. Buyers would not sue in the case of high/disagree because they knew with certainty that the verifier had tested (a disagree report could be issued only if the verifier tested). In the case of low/agree, the buyers know the asset is low because sellers could not disclose high assets as low; although buyers could still sue in this case, the expected value from doing so was negative. That is, buyers can expect to receive 0.10 ($4)+0.90 (P−$0.40), implying a breakeven point of $0.40.

[24] This is the British system of incorporating legal costs. The effect of different methods of allocating legal costs is discussed in King and Wallin (1990). In security cases in the United States, courts can pass on all costs to the losing parties. The British system was selected because it should be easier for subjects to understand. The system produces predicted strategies that do not differ, qualitatively, from the American system.

Step 6 would also apply under the strict liability regime but with these modifications. A suit was automatically filed against a verifier given the above four conditions, and that verifier had to pay the damages plus the enforcement cost whether or not T_Y had been selected.[26] The maximum amount of damages that could be assessed on the verifiers in any one suit under both the negligence and strict liability regimes was fixed at \$3.60 (i.e., \$4 − \$0.40).[27]

III. Hypotheses

In this section we develop predictions for strategies of verifiers, sellers, and buyers in each regime. Comparative hypotheses based on these strategies are then provided.

Predictions for the No-liability Markets

We will begin by providing predictions of buyers' strategies. We assume that buyers are risk-neutral[28] and bid their reservation value (consistent with the second-price, sealed-bid institution) with the expectation of zero expected profits.

If buyers actually know the asset is of high-quality, they are predicted to price the asset at its expected value of \$3.60 (90 percent of \$4). Similarly, if buyers actually know the asset is of low-quality they are predicted to price it at \$0.40 (10 percent of \$4). The expected return to buyers if sellers selected the cost-free (costly) investment is \$1.04 (\$2.96).[29] Buyers never learn the investment choice, and they will know the asset quality only if sellers disclose an asset as low or if a verifier disagrees with a seller's high disclosure after testing. In this regime (for reasons given below) sellers are not expected to select the costly investment, and verifiers, if hired, are not expected to test.

Sellers will hire verifiers only if this would increase their expected net profits, which will be greater if verifiers provide a credible report that supports the seller's decision to select I_Y. A credible report will be provided if the verifiers test to determine the asset quality. But the test decision by verifiers is made after the sellers have committed to paying their offer fees. Thus, the verifiers have the choice of either not testing and retaining the offer fees (which requires them to send the agree report) or paying

---

[25] The amount of damages assessed against a verifier who selected T_N was the amount required to reimburse buyers for the excess over what they would have paid if they had known the asset was a q_L type. We assume that buyers would value a q_L asset at its expected value of \$0.40. Note that a verifier who is sued is assessed damages if T_N was selected even with a q_H asset.

[26] The fact that suits were triggered automatically in the strict liability regime created a possibility of suits when the high/disagree combination occurred even though buyers bid \$0.40 (or lower). Buyers would not be reimbursed in these cases, but verifiers would still have to pay enforcement costs. We circumvented this problem by imposing a minimum of \$0.55 before a suit was automatically triggered. Thus any risk-seeking buyers who priced the asset above \$0.40 but below \$0.55 would not force the verifier to bear the \$1 enforcement cost. We also imposed the same minimum in the negligence regime to maintain consistency, but it did not produce any qualitative effects on buyers' bidding strategies. See footnote 23 above.

[27] Although the strict liability regime is rather extreme, it serves our purposes in that it provides a high legal standard compared with the negligence liability regime.

[28] The assumption of risk-neutrality does not qualitatively affect the hypotheses.

[29] The expected values are calculated as $(0.90 \times 0.20 + 0.10 \times 0.80)\$4$ and $(0.90 \times 0.80 + 0.10 \times 0.20)\$4$.

$0.50 for the test that informs them of the actual asset quality, which in turn generates a truthful report. In a single period world, once hired, the self-interested verifiers will not test, will retain their offer fee, and will report agree by default.

The strategy for a verifier in a multiperiod game could differ from that of a single period game. For example, a verifier might attempt to form a "reputation" by testing in early periods in anticipation of continuing to be hired in the remaining periods of the market. There are at least two reasons why a verifier will have difficulty forming such a reputation. The first is based on the notion of subgame perfection. In the last period, the verifier (if hired) would not test, as predicted above. Subgame perfection implies that, given the expected strategy choice in the last period, the second to last must have an identical prediction because all participants know what will happen in the last period. This unraveling unfolds back to the first period. Hence, reputation formation would not be possible because the verifiers would never be hired, or, if hired in period t, they would not test since they would expect not to be hired in period $t+1$.

The second reason that reputations may not be formed by verifiers stems from a potential independence problem. To form a reputation, the verifier must test. But if a verifier tests and then disagrees with the seller's disclosure of high because the asset quality is actually low, the seller may respond by not hiring that particular verifier in subsequent periods. It is sufficient that the threat of this behavior exists. For these two reasons, we predict that, if hired, the verifier will not test in the no-liability regime. We also predict that verifiers in the no-liability regime will compete with one another to be hired, continuously lowering their offers to the point that they converge toward zero (the cost of not testing).

We base our predictions of the sellers' strategies on the assumption that sellers capture all excess surplus. The reason is that buyers bid and verifiers offer (in a Bertrand-type competition) so as to compete away all economic surplus. The expected profits to the sellers are calculated below for the four possible verification hiring/investment conditions of V_N/I_N, V_N/I_Y, V_Y/I_Y, and V_Y/I_N. The V_N/I_N combination is predicted to be the only stable equilibrium in our finite setting for the no-liability markets.

Sellers' Expected Profits from Selecting V_N/I_N. Sellers can expect to receive an average price of $1.04 for the asset if the buyers price the asset as if the seller always selected I_N when a verifier is not hired (selection of V_N and I_N result in zero costs to the sellers).

Sellers' Expected Profits from Selecting V_N/I_Y. If the seller selects the V_N/I_Y combination and the buyers price the asset accordingly, the seller's net expected profits would be $2.21 ($2.96 − $0.75). This outcome requires that both the seller and the buyers cooperate.[30] The critical requirement for the V_N/I_Y strategy to be an equilibrium is that the buyers believe the seller will continue to consistently select I_Y. A large body of experimental research has shown that strategic agents may cooperate if they believe that their opponents will cooperate (Axelrod 1984). The solution concept of subgame perfection in a finite horizon setting, however, predicts that such cooperative behavior is not stable.

Sellers' Expected Profits from Selecting V_Y/I_Y. If a seller selects both V_Y and I_Y, the expected net return would be $2.96 minus $0.75 for the cost of I_Y and less the verifier's

---

[30] Cooperation is used here to refer to a strategy combination in which the seller consistently selects V_N and I_Y and the buyers bid as if I_Y were selected.

Table 2

Summary of the Predictions of Strategies for Sellers, Verifiers, and Buyers for the Three Liability Regimes

	No Legal Liability			
	Single Period	*Multiperiod*	*Negligence Liability*	*Strict Liability*
Sellers				
Percentage of verifier hires (V_Y)	0	>0	100	100
Percentage of costly investment (I_Y)	0	>0	100	100
Percentage of false disclosures (θ)	100	<100	100	100
Verifiers				
Offers for verification services	$0	>$0	$0.50	$0.86
Percentage of tests (T_Y)	0	>0	91	100
Buyers				
Price, given high/agree message	$1.04	>$1.04	$3.53	$3.96
Buyers' suing strategy (λ)	N.A.	N.A.	47%	N.A.

fee. But the self-interested verifier will select T_N, retain the fee, and report agree by default. The seller anticipates the verifier's strategy and assumes that buyers will also. Hence, the seller cannot expect a price of $2.96 because the buyers will not rely on the seller's disclosure and the verifier's report.

Sellers' Expected Profits from Selecting V_Y/I_N. The last possibility is for the seller to select the V_Y/I_N combination, which will generate an expected price of $1.04 minus the verifier's fee. This is less attractive than the V_N/I_N combination.

Summarizing, the predictions for a single-period setting for the no-liability regime are that (1) the verifiers will make offers that will converge to zero over time and, if hired, will not test; (2) sellers will anticipate the verifiers' strategy and will not hire them and will not make the costly investment; and (3) buyers will ignore the sellers' disclosures (and any verifier reports) and price the assets at $1.04. In a multiperiod setting, these predictions might have to be modified to recognize the possibility that a seller may invest more frequently, may disclose the low-quality assets truthfully, and may hire a verifier (who tests in order to generate credibility), at least in earlier periods of the markets. Table 2 shows the predictions for this and the other two regimes.

Predictions for the Negligence Liability Markets

Under the negligence liability regime, the buyer who purchased the asset could sue to determine whether the verifier tested, provided that (1) a verifier was hired, (2) the seller disclosed the asset as high, (3) the zero payoff was obtained, and (4) the purchasing buyer paid more than $0.55. Implicit in these conditions is that the verifier reported agree, given that the seller disclosed high. That is, buyers would not sue if the verifier reported disagree since they would know that the verifier could not issue such a report without first selecting the verification test. Recall that in the negligence regime a verifier who tested would not have to pay damages to the purchasing buyer. Hence, the decision to sue was rational only if the buyer received a high disclosure coupled with an agree verification report.

A Nash equilibrium solution can be obtained by solving three equations, one each for the buyers' suing strategy, the buyers' pricing strategy, and the verifier's testing strategy. These three equilibrium strategies depend upon the investment and disclosure strategies of the seller. The solutions we provide are for the case in which the seller makes the costly investment. Since the structure of the other solutions is similar, they are omitted to conserve space. The details supporting these equations are provided by the decision tree in figure 1.

The buyers' equilibrium suing strategy is obtained by equating the expected benefits from suing with the expected costs from not suing, which is:

$$(P^*-0.40)\frac{(0.08+0.18\Theta)\Omega}{(0.08+0.18\Theta\Omega)}-\left[1-\frac{(0.08+0.18\Theta)\Omega}{(0.08+0.18\Theta\Omega)}\right]\$1=0, \tag{1}$$

where Θ represents the probability of the seller's disclosing a q_L asset as high and Ω is the probability of the verifier's not testing. The first component on the left hand side of the equation represents the damages the buyer receives (from the verifier if the buyer sues and the verifier shirked) multiplied by the probability the verifier shirked and the seller misrepresented the asset quality. As the suing information set from figure 1 shows, the probability is calculated from the three possible cases facing the buyer: (1) the seller had a q_H asset, and the verifier tested and justifiably reported agree, (2) the seller had a q_H asset, but the verifier did not test and reported agree by default, and (3) the seller had a q_L asset, but the verifier did not test and again reported agree by default.[31] The second component on the left hand side is the probability the verifier tested multiplied by the $1 enforcement cost the buyer must pay when a suit is brought and the verifier had tested.

The equilibrium price is determined by solving for the probability that the seller's asset has a $4 payoff multiplied by $4. The probability of a q_H asset is a function of the verifier's testing strategy and the seller's disclosure strategy. The equilibrium price is:[32]

$$P^*=\frac{2.88+0.08\Theta\Omega}{0.8+0.2\Theta\Omega}. \tag{2}$$

---

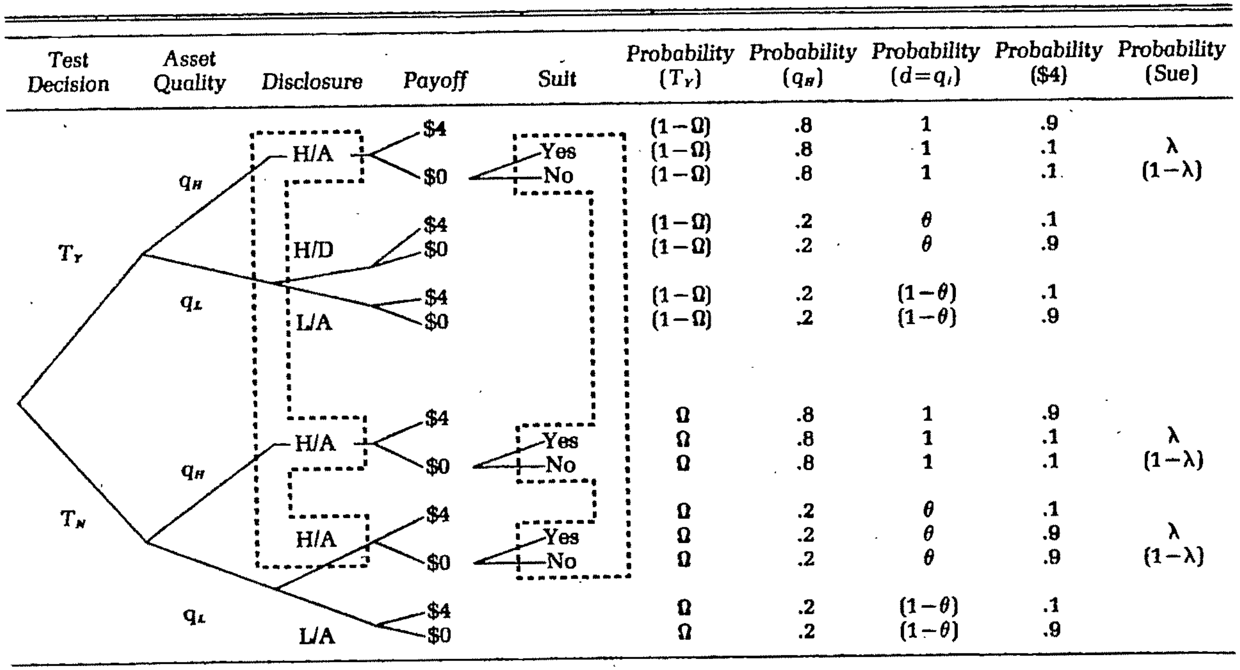
[31] The probability of the buyer's winning the suit is the probability of the verifier's not testing, given that a verifier was hired, that the message was high/agree, and that the asset's payoff was zero (shown in figure 1 as the bottom two nodes of the suing information set) divided by the probability that a verifier was hired, that the message was high/agree, and that the asset's payoff was zero (shown in figure 1 as all three nodes of the suing information set).

[32] The price is the expected cash payoffs from the assets (shown in figure 1 as the pricing information set, which is those nodes with the H/A message).

Figure 1

Game Tree for Negligence Liability Regime

(When Seller Selects Costly Investment and Verifier is Hired)


Test Decision
Asset Quality
Disclosure
Payoff
Suit
Probability (T_Y)
Probability (q_H)
Probability ($d=q_i$)
Probability ($4)
Probability (Sue)
T_Y
T_N
q_H
q_L
q_H
q_L
H/A
H/D
L/A
H/A
H/A
L/A
$4
$0
$4
$0
$4
$0
$4
$0
$4
$0
$4
$0
Yes
No
Yes
No
Yes
No
$(1-\Omega)$ .8 1 .9
$(1-\Omega)$ .8 1 .1 λ
$(1-\Omega)$ .8 1 .1 $(1-\lambda)$
$(1-\Omega)$ .2 θ .1
$(1-\Omega)$ .2 θ .9
$(1-\Omega)$ .2 $(1-\theta)$ .1
$(1-\Omega)$ .2 $(1-\theta)$ .9
Ω .8 1 .9
Ω .8 1 .1 λ
Ω .8 1 .1 $(1-\lambda)$
Ω .2 θ .1
Ω .2 θ .9 λ
Ω .2 θ .9 $(1-\lambda)$
Ω .2 $(1-\theta)$ .1
Ω .2 $(1-\theta)$ .9

T_Y = Verifier tested the asset.
T_N = Verifier did not test the asset.
q_H = High-quality asset.
q_L = Low-quality asset.
$d=q_i$ = Seller disclosed truthfully.
H/A = Seller disclosed asset as high and verifier agreed.
H/D = Seller disclosed asset as high and verifier disagreed.
L/A = Seller disclosed asset as low and verifier agreed.
Ω = Probability verifier tested.
θ = Probability seller disclosed a q_L asset as high.
λ = Probability purchasing buyer sued.

Finally, the verifier's testing strategy is selected to make the buyers indifferent to suing and can be calculated by setting the costs of testing ($0.50) equal to the expected cost of losing a suit, or:

$$0.5=\lambda(0.08+0.18\Theta)(P^*+0.6), \qquad (3)$$

where λ is the probability of the purchasing buyer's suing. By solving the three equations (assuming the seller always misrepresents q_L assets as high), the equilibrium strategies when the costly investment is selected is for the verifier to shirk (select T_N) 8.9 percent of the time and the buyers to sue 47 percent of the time, which translates into a price of $3.53 (given a message of high/agree). This equilibrium has the verifiers offering their services at $0.50 and sellers generating a net profit of $1.71.[33] If the cost-free investment is made, the equilibrium testing strategy is to shirk less than 1 percent of the time and the suing strategy to sue about 16.5 percent of the time, which translates into a price (given a high/agree combination) of $3.49. This equilibrium has the verifiers offering their services at $0.50 and sellers generating a net profit of $0.54. Since this is less than the expected profit of $1.71 shown above, sellers are predicted to make the costly investment.

Predictions for the Strict Liability Markets

In the strict liability markets, suits were automatically invoked ($\lambda=1$) whenever a seller disclosed the asset as high, a verifier was hired, the zero payoff obtained, and the price paid for the asset was greater than $0.55. In these markets the verifier was always responsible for damages of ($P^*-0.40$) and the enforcement cost of $1. Using the same approach as in the negligence liability regime above, we obtain the following predictions if the seller selects the costly investment: verifiers will always test, and the buyers will price the asset at $3.96 (given a message of high/agree). The verifiers' equilibrium offers will be $0.86[34] and the sellers' net profits will be $1.63. The $0.86 amount includes $0.50 for the cost of testing and $0.36 for the expected costs of reimbursing the buyers and paying the enforcement costs.

If the seller selects the cost-free investment instead, verifiers will again always test and the buyers will price the asset at $3.96 (given a high/agree message). However, because of the higher frequency of high/disagree reports, the verifiers' equilibrium offers then drop to $0.59 and the sellers' net profits drop to $0.52. Since this is lower than the expected profits of $1.63 above, the seller has incentives to hire credible verifiers and to then select the costly investment.

Hypotheses

The above predictions are the basis for developing the following five testable hypotheses that compare the market outcomes in the three liability regimes. Hypothesis H6 compares the regimes in terms of overall efficiency, given the first five hypotheses. The hypotheses are specified as comparisons across regimes rather than as tests of point predictions because our main interest is to investigate relative market outcomes over the three regimes.

---

[33] The seller's net proceeds when the costly investment is made is $0.8P^*+0.2[0.4-0.4\Theta\Omega+\Theta\Omega P^*]-0.75$ $-$the verifier's fee, and is $0.2P^*+0.8[0.4-0.4\Theta\Omega+\Theta\Omega P^*]-$verifier's fee when the cost-free investment is made.

[34] The verifier's offer when the seller is expected to select the cost-free investment equals $0.02P^*+0.012$ $+\Theta\Omega[0.72P^*+0.432]+0.5-0.5\Omega$, and equals $0.08P^*+0.048+\Theta\Omega[0.18P^*+0.108]+0.5-0.5\Omega$ if the seller is expected to select the costly investment.

Size of the verifiers' offers

H1: The offers will be highest in the strict liability markets, second highest in the negligence, and lowest in the no-liability markets.

Frequency of sellers' hiring a verifier

H2: The frequency of hiring verifiers will be the lowest in the no-liability markets and approximately equal in the negligence and strict liability markets.

Frequency of sellers selecting the costly investment

H3: The frequency of the costly investments will be lowest in the no-liability markets and approximately equal in the negligence and strict liability markets.

Verifiers' testing decision

H4: Verifiers will test with the highest frequency in the strict liability markets, with the second highest frequency in the negligence liability markets, and with the lowest in the no-liability markets.

Pricing of the asset by the buyers

H5: In periods when sellers disclose high and the verifiers report agree, buyers will price the asset the highest in the strict liability markets, second highest in the negligence liability markets, and lowest in the no-liability markets.[35]

Efficiency of the three regimes

H6: Given hypotheses H2, H3, and H4, the levels of efficiency will be approximately equal in the strict and negligence markets, and lowest in the no-liability markets.

IV. Results

The results from all of the markets in each liability setting appear in appendix A (available from the authors), and are reported separately for two time segments: periods 1–15 and periods 16–20. The latter are shown separately to indicate whether there are end-period effects. Although there were some changes between periods in the no-liability markets, no changes were detected that would alter our inferences. Hence, we base our analysis on observations from all 20 periods in each market.

Hypothesis H1 (Size of offers)

The results reported in table 3 (top row) are consistent with this hypothesis. The offers in the no-liability markets averaged $0.39 over all 20 periods ($0.44 and $0.23 for the two time segments, respectively), compared with $0.75 in the negligence markets ($0.76 and $0.71 for the two time segments) and to $1.43 in the strict liability markets

---

[35] Our price predictions for cases in which sellers disclose low or verifier disagree is $0.40 for all regimes. When sellers disclose high and no verifier was hired, the predicted price is $1.04 for the three regimes.

Table 3
Summary of Market Data

		No Legal Liability			*Negligence Liability*			*Strict Liability*		
		Periods			*Periods*			*Periods*		
Row		*1–15*	*16–20*	*1–20*	*1–15*	*16–20*	*1–20*	*1–15*	*16–20*	*1–20*
1	Verification Offers									
	(Mean)	$0.44	$0.23	$0.39	$0.76	$0.71	$0.75	$1.41	$1.47	$1.43
	(Standard Deviation)			$0.24			$0.30			$0.92
	(Median)			$0.40			$0.70			$1.08
2	Accepted Offers									
	(Mean)	$0.34	$0.13	$0.28	$0.75	$0.66	$0.70	$0.83	$0.63	$0.78
	(Standard Deviation)			$0.20			$0.08			$0.45
	(Median)			$0.25			$0.70			$0.60
3	Number of Verifier Hires (V_Y)	52	21	73	136	50	186	86	29	115
4	Number of V_Ys divided by number of periods*	58%	70%	61%	76%	83%	78%	48%	48%	48%
5	Number of costly investments (I_Y)	46	8	54	141	48	189	122	41	163
6	Number of I_Ys divided by number of periods	51%	27%	45%	78%	80%	79%	68%	68%	68%
7	Number of I_Ys given V_Y	26	5	31	116	48	164	70	28	98
8	Number of tests (T_Y)	6	0	6	100	39	139	59	21	80
9	Number of T_Ys divided by the number of V_Ys	12%	0%	8%	74%	78%	75%	69%	72%	70%

* The number of seller periods in the no-liability markets was 90, 30, and 120 for the intervals of 1–15, 16–20, and 1–20, respectively, and 180, 60, and 240 for the same intervals in the negligence and strict liability markets.

($1.41 and $1.47 for the two time segments).[36] The differences between the no-liability and the negligence liability markets, the no-liability and strict liability markets, and the negligence and strict liability markets are all significant ($p<0.01$) which is consistent with hypothesis H1.[37]

Hypothesis H1 is also supported by the average accepted offers (second row of table 3). The average accepted offer of $0.28 in the no-liability markets was significantly lower ($p<0.01$) than the average accepted offers of $0.70 and $0.78 in the two liability regimes, but the difference between the latter two averages was less significant ($p<0.05$).

---

[36] There were 240 offers in the no-liability markets (three markets in which two verifiers made two offers to each of two sellers for 20 periods), and 480 offers in both the negligence and strict markets ($6\times2\times2\times20$).

[37] Statistical analysis was performed with both parametric tests and approximate randomization. The results were qualitatively the same with the different approaches. The approximate randomization calculations are computed by generating a test statistic from the actual data, then the data are combined and divided repeatedly and the statistic is computed for each of the random divisions. The p-values reported represent the percentage of times the statistic from the random divisions is less than the test statistic for the actual data. All analyses were done with 40,000 repetitions. See Edgington (1987) for a discussion of the approximate randomization method.

Hypothesis H2 (Hiring of verifiers)

The results reported in table 3 (third and fourth rows) do not support this hypothesis. Verifiers were hired in 61 percent of the periods in the no-liability markets (73/120), compared with 78 percent of the periods in the negligence markets (186/240) and 48 percent of the periods in the strict liability markets (115/240). The difference in percentages of hires between the negligence and the strict liability markets is significant ($p<0.05$).[38] However, the differences between the percentage of hires in the no-liability markets and those in the negligence and strict liability markets are not significant.

The results for the individual markets shown in appendix A indicate that verifiers were hired only eight out of 80 opportunities by the four sellers in the first two strict liability markets. Verifiers were seldom hired mainly because their offers in these markets averaged more than $2. In contrast, verifiers were hired more frequently in the other strict liability markets when their average offers were below $1. This is an indication of the difficulties that verifiers had in pricing their services in the strict liability markets. As shown in appendix A verifiers' offers and their subsequent hires were much more evenly distributed across the negligence markets.

Hypothesis H3 (Frequency of costly investment)

The number and percentage of I_Y selections by the sellers appear in the fifth and sixth rows of table 3. These results are partially supportive of hypothesis H3 since the difference between the 54 selections of I_Y in the no-liability markets (45 percent) and the 189 selections in the negligence markets (79 percent) is significant ($p<0.01$), and the difference between the negligence and the strict liability markets (163 or 68 percent) is not significant, as predicted. However, the difference between the percentages of costly investments in the no-liability and the strict liability markets (45 percent and 68 percent, respectively) is also not significant, contrary to the predictions. The seventh row in table 3 shows that the seller selected the costly investment in 31 of the 73 periods in which a verifier was hired in the no-liability markets (42 percent), compared with 164 of the 186 periods for the negligence markets (88 percent) and 98 of the 115 periods for the strict markets (85 percent). This indicates that sellers were more likely to make the costly investment when they also hired a credible verifier.

Hypothesis H4 (Verifier's testing decision)

The results in rows 8 and 9 of table 3 are partially consistent with hypothesis H4 in that the verifiers tested least often in the no-liability markets—six of 73 periods hired (8 percent) versus 139 of 186 periods in the negligence markets (75 percent) and 80 of 115 periods in the strict liability markets (70 percent). The differences between the percentages of testing in the no-liability markets and those in the negligence and strict markets are significant ($p<0.01$). However, verifiers did not test significantly more often in the strict liability markets than in the negligence liability markets. The latter result suggests that the imposition of a strict liability rule in the market settings did not lead to an increase in the quality of verifier's services. This may again be reflective of the problems the verifiers encountered in pricing their twofold service of insurance and verification in the strict liability setting.

---

[38] This analysis used each seller as the unit of analysis. Thus, there were six observations in the no-liability markets and 12 in each of the other two regimes.

Table 4
Pricing Based on Buyers' Information

	No Legal Liability Markets	*Negligence Liability Markets*	*Strict Liability Markets*
Panel A			
Seller disclosed high and verifier agreed			
Price			
Mean	$2.15	$3.28	$3.50
Standard Deviation	$0.85	$0.42	$0.56
Median	$2.07	$3.35	$3.75
Frequency	68	147	94
Conditional Value*	$2.09	$3.36	$3.28
Panel B			
Seller disclosed low, or verifier disagreed			
Price			
Mean	$0.32	$0.38	$0.47
Standard Deviation	$0.07	$0.08	$0.14
Median	$0.29	$0.40	$0.47
Frequency	7	44	31
Conditional Value	$0.40	$0.40	$0.40
Panel C			
Seller disclosed high and a verifier was not hired			
Price			
Mean	$1.84	$1.55	$2.18
Standard Deviation	$0.72	$0.81	$1.09
Median	$1.60	$1.50	$2.25
Frequency	45	49	115
Conditional Value	$2.32	$1.96	$2.20

* The conditional value is based on the relative number of high and low assets.

Hypothesis H5 (Pricing by buyers)

The results pertaining to this hypothesis appear in panel A of table 4, and are consistent with hypothesis H5. Buyers paid a mean (median) price of $3.50 ($3.75) over the 94 occurrences of the high/agree message in the strict liability markets, compared to mean (median) prices of $3.28 ($3.35) over the 147 occurrences in the negligence liability markets and $2.15 ($2.07) over the 68 occurrences in the no-liability markets.

The prices in the no-liability markets were significantly lower than those in the negligence and strict liability markets ($p<0.01$), and the prices in the negligent liability markets were significantly lower than those in the strict liability markets ($p<0.02$). The results are consistent with the expectation that buyers in the no-liability markets did not regard the verifiers' reports as credible, and that they viewed the verifiers in the strict liability regime as providers of a form of insurance. Also notice that the standard deviation of $0.42 was lower in the negligence liability markets than in the other two regimes, which is consistent with the notion that the verifiers' reports were viewed as being more credible in those markets.

The last line of panel A shows the conditional value of assets, which is the expected value computed by using the actual numbers of high- and low-quality assets the sellers had for sale. These are provided to ascertain whether the buyers' pricing strategies were consistent with the actual asset quality held by the sellers. The data indicate that the conditional values in the no-liability and the negligence liability markets were approximately equal to the median prices, and fairly close to the mean prices. In contrast, the conditional value in the strict liability markets was below the mean and median prices actually paid. Again, this suggests that verifiers in those markets were viewed as providing insurance to the buyers.

The remaining data in table 4 are provided to indicate whether buyers' pricing strategies were consistent with the information they received in two other situations. Panel B provides the mean and median prices for the situation in which buyers knew the sellers had a q_L asset for sale (either sellers disclosed the asset as low or verifiers disagreed with sellers' high disclosures). Although we predicted that buyers would price the assets in this situation at $0.40 (0.10 × $4) across all three regimes, they actually paid an average of $0.32 (standard deviation of $0.07) and a median price of $0.29 for the seven occurrences in the no-liability markets, an average of $0.38 (standard deviation of $0.08) and a median price of $0.40 for the 44 occurrences in the negligence liability markets, and an average price of $0.47 (standard deviation of $0.14) and a median price of $0.47 for the 31 occurrences in the strict liability markets.

Asset prices might have deviated from the predicted price of $0.40 in the no-liability and strict liability markets for several reasons, one being that buyers were not risk-neutral.[39] Another is that the second-price sealed-bid institution did not provide incentives for buyers to bid their reservation value. A third is that the multiple-period aspect of the markets encouraged buyers to "speculate" on runs in the state payoffs. Finally, buyers might have been influenced by the insurance function provided by verifiers in the strict liability markets (in six periods, buyers paid $0.55 or more when the combination was high/disagree, which triggered an automatic lawsuit if the asset had the zero payoff).

Panel C shows prices for the situation in which sellers disclosed high but no verifier was hired. Recall that buyers were predicted to price assets in this situation at $1.04 in all three regimes, assuming the sellers always selected the cost-free investment and disclosed all assets as high. Actual prices might have deviated from this extreme prediction if sellers tried to establish reputations for selecting the costly investment or for truthfully revealing low assets. The conditional values indicate that sellers in all three regimes did not consistently select the cost-free investment since the values are much higher than the predicted price of $1.04 in all three cases. However, buyers did not completely reward the sellers for selecting the costly investment except in the strict liability markets (conditional value of $2.20 vs. the actual mean price of $2.18). Note that actual prices and conditional values in the two liability regimes in this situation still fell

---

[39] In five of the seven periods in the no-liability markets, buyers paid less than $0.40 for the low asset, and they paid $0.41 in the other two periods; in contrast, buyers in a strict liability market paid $0.47 or more six times.

Table 5
Efficiency Levels for the Three Regimes

	No Legal Liability Markets	*Negligence Liability Markets*	*Strict Liability Markets*
Optimal economic resources (cents per seller per period)	$2.21	$2.21	$2.21
Less:			
Lawsuits[a]	−$0.00	−$0.10	−$0.08
Costs from selecting I_N[b]	−$0.64	−$0.25	−$0.38
Resources not considering verification costs	$1.57	$1.86	$1.75
(% of Maximum)	(71%)	(84%)	(79%)
Cost of verification tests	−$0.03	−$0.29	−$0.17
Net economy resources	$1.54	$1.57	$1.58
(% of Maximum)	(70%)	(71%)	(71%)

[a] Calculated as $1 multiplied by the frequency of suits divided by the number of periods.

[b] Calculated as ($2.21 − $1.04) times the frequency of the cost-free investments divided by the number of periods.

significantly below those shown in panel A when verifiers' reports supplemented sellers' high disclosures ($p<0.01$).

Taken overall, buyers' pricing strategies appear to have been consistent with the information they received in the three different pricing situations. Their pricing strategies also appear to have been influenced by the presence or absence of a credible verifier, as in the two liability settings.

Hypothesis H6 (Market efficiency)

The efficiency measures used to evaluate hypothesis H6 consist of the average amounts of economic resources generated in the three liability settings. The calculations are shown in table 5.

The top line in the table shows the expected economic wealth that could be generated by each seller each period if the costly investment were made, which is $2.21 (0.74 × $4 − $0.75). We then subtract the average costs of suits, if applicable, and the average (opportunity) cost of sellers' selecting the cost-free investment. The average lawsuit cost per period is about equal for the two liability settings (there were 25 and 20 suits, respectively, in the negligence and strict liability settings).

The results indicate that the efficiency level was significantly higher in the negligence and strict liability markets than in the no-liability markets ($p<0.01$), and significantly higher in the negligence markets than in the strict liability markets ($p<0.02$). After subtracting the cost of testing, however, the efficiency levels do not differ among the three regimes, primarily because more verifiers were hired and then chose to test in the negligence setting. This provides mixed support for hypothesis H6.

The verifier's cost of testing, $0.50, is an arbitrary figure and somewhat unrealistic when compared with the expected value of the asset ($2.96) and the cost of the investment ($0.75). For example, the average audit fees for large clients typically is only a fraction of 1 percent of sales (or assets). We performed some sensitivity analysis on our data, which showed that, had testing cost $0.14 or less, the negligence liability regime would have been significantly more efficient than the no-liability regime even after considering the cost of testing, *ceteris paribus*. To make a significant difference between the negligence and strict liability markets the cost of testing would have had to be about $0.01.

The cash payments to the participants were analyzed also to determine if these systematically differed across the regimes. We found no significant difference across the regimes for all three roles. However, there is a fundamental problem inherent in using cash payments as a basis for comparison of efficiency because the random draws for the asset qualities and the payoff amounts could differ in markets even though the investment and testing strategies did not.

Finally, there could be some secondary effects of the verification service that are not captured in our efficiency calculations. As one example, a credible verification report could lead to a higher consensus among the buyers of the value of the sellers' assets. The fact that the standard deviation of prices shown in table 4 for the high/agree combination is significantly lower in the negligence and strict liability settings than in the no-liability regime is consistent with this possibility.

V. Summary and Conclusions

Overall, the present results (and those obtained in a previous study [Dopuch and King 1991]) suggest several conclusions. First, as in most previous experimental economic studies, economic agents appear to adopt complex strategies, including attempts to form "reputations" over multiperiod settings. These strategies are influenced by the institutions that coordinate the interactions among agents. Second, we find that the institution of an *ex post* legal system can influence the demand for verification services. Specifically, both the negligence and the strict liability regimes provided incentives to the verifiers to balance the costs of providing their services with the expected (litigation) cost of not providing the services. With credible verification, sellers were motivated to acquire the verification services voluntarily and to make investment decisions that led to higher levels of efficiency than when no such credible reports were available (as in the no-liability setting).[40] At the same time, we found no evidence indicating that the imposition of a strict liability rule on the verification service would provide any net benefits beyond a negligence rule.

Lastly, we find the experimental approach appears to be a useful tool in aiding our understanding of the functioning of auditing markets. Using this method in combination with theoretical and empirical work from naturally occurring markets can provide insights into the operations of auditing markets. Such insights may give policy makers cause to ponder proposed changes in the present institutional framework of auditing, such as expanding auditors' legal liability.

---

[40] Our market setting, which allowed the demand for auditing services to arise endogenously under a negligence liability rule, is similar to the market for auditing services that existed in the United States in the early part of this century. The fact that voluntary hires took place in almost 80 percent of these market periods is consistent with the Watts and Zimmerman (1983) finding that many firms voluntarily had audits during that time period.

Extensions of our study could focus on various conditions that might further affect the relative net benefits of strict versus negligence regimes (e.g., introducing alpha and beta errors in the enforcement system under the negligence rule). Another possibility would be to allow sellers and verifiers to agree on a total disclosure package for the buyers (which might lead to more conservative messages). The latter alteration could also allow an investigation of the effects of imposing joint and several liability rules on sellers and verifiers.

POSTSCRIPT*

We are pleased that in the final section of their "Perspective," DeJong and Forsythe offer an "alternative research program" that parallels a research agenda we have been attempting to follow during the past few years. This can be illustrated by summarizing a sequence of related studies that we, along with Wallin, have conducted and that reflects this research agenda.

1. Early precedents for our work in auditing are two studies by King and Wallin (*Contemporary Accounting Research*, Spring 1990; and *Journal of Accounting Research*, Spring 1991), who tested theories of voluntary disclosures by sellers of assets in a market setting involving only sellers and buyers. Their setting did not incorporate legal liability penalties for false disclosures by the sellers, although an anti-fraud rule was imposed on sellers' disclosures in some of their markets.
2. The basic King/Wallin setting was modified by Dopuch, King, and Wallin (*Auditing: A Journal of Practice & Theory*, Supplement 1989) to include a computerized (perfect) verification mechanism that could add credibility to the sellers' voluntary disclosures. The computerized mechanism could be purchased by the seller or by one or more buyers, with the resulting report provided in either case to all buyers. The verification mechanism had an endogenous value since it encouraged the sellers to make costly investments that improved the welfare of their experimental economy.
3. Dopuch and King (Canadian Certified General Accountants *Research Monograph* #19, 1991) extended this work along two dimensions: adding a human verifier who was subject to a moral hazard problem, and making the verification report private information to the buyer(s) who actually purchased the verification service. Neither the sellers nor the human verifiers were subject to legal penalties for lying or shirking.
4. Our liability study, which appears as part of this Forum, extends the work in (3) by imposing alternative liability rules on the verifier. The reader will note that in the other paper in this Forum, Wallin extends the work referred to in (2) by imposing a liability rule on the sellers.
5. In our most recent study (Dopuch and King, *Journal of Accounting Research*, Supplement 1991), we investigated potential effects on auditor independence and pricing of services when the verifiers are allowed to offer two interdependent services to the same sellers—a management advisory service (MAS) and a verification service. In this setting, the verifier was subject to a negligence liability rule for the verification service but not for the MAS service.

> This study made us particularly sensitive to the complexity issue stressed by DeJong and Forsythe since we observed our subjects encountering considerable difficulties in understanding their tasks which we were able to overcome by allowing them to work as "teams."

In summary, we fully agree with the basic thrust of the research program proposed by DeJong and Forsythe; the lack of theories that incorporate even simple institutional arrangements in auditing forces us to proceed a step at a time. Each research, however, should be encouraged to define his or her own sequence of steps in developing alternative research agendas in auditing.

References

Adler, S. 1989. Can juries do justice to complex suits? *The Wall Street Journal* (December 21): B1.

Axelrod, R. 1984. *The Evolution of Cooperation.* New York: Basic Books.

Cave, G. 1985. Extensions of accountants' liability for negligence: One step closer to a new implied warranty of results. *University of Colorado Law Review* 56 (Winter): 265–87.

Chilton, K. 1990. *Product Liability Reform: Debating the Issue.* Center for the Study of American Business, Washington University.

Cox, J., B. Roberson, and V. Smith. 1982. Theory and behavior of single object auctions. In *Research in Experimental Economics,* Vol. 2. Greenwich, CT: JAI. 1–43.

Cowan, A. 1990. Rivalries, responsibilities and some new risks. *The New York Times* (March 11): F11.

Davis, J. 1989. Auditor bidding behavior and independence: A laboratory markets investigation. Doctoral dissertation, University of Arizona.

DeJong, D., and J. Smith. 1984. The determination of audit responsibilities: An application of agency theory. *Auditing: A Journal of Practice & Theory* 4 (Fall): 20–34.

———, R. Forsythe, and W. Uecker. 1985. The methodology of laboratory markets and its implications from agency research in accounting and auditing. *Journal of Accounting Research* 23 (Autumn): 753–93.

Dopuch, N., and R. King. 1991. Experimental tests of auditing as a credibility generating mechanism. *Canadian Certified General Accountants' Research Foundation Monograph* (No. 19).

Edgington, E. 1987. *Randomization Tests.* 2nd ed. New York: Marcell Dekker.

Forsythe, R., and R. Lundholm. 1990. Information aggregation in an experimental market. *Econometrica* 58 (March): 309–47.

Gormley, R. 1985. The foreseen, the foreseeable, and beyond—Accountants' liability to nonclients. *Defense Law Journal* 34 (February): 75–123.

Guy, D., C. Alderman, and A. Winters. 1990. *Auditing.* New York: Harcourt Brace Jovanovich.

Kachelmeier, S. 1990. A laboratory market investigation of the demand for auditing in an environment of moral hazard. *Auditing, A Journal of Practice & Theory* (forthcoming).

King, R., and D. Wallin. 1990. Legal cost allocation and incentives to issue fraudulent disclosures: An analytical assessment. Working paper, The Ohio State University.

Kinsella, J. 1984. Protecting the auditor from unwarranted third-party liability: Rethinking the indemnification issues. *Syracuse Law Review* 35 (No. 2): 763–91.

**Editor's Note:* Instead of offering the pages of *The Accounting Review* to debate the similarities and differences between the two research programs (the one proposed by DeJong and Forsythe and that adopted by Dopuch and King), I have agreed to append this postscript to the Dopuch and King article. Readers interested in this area should be able to compare those two programs. DeJong and Forsythe have acceded to this option.

Kothari, S. P., T. Lys, C. Smith, and R. Watts. 1988. Auditor liability and information disclosure. *Journal of Accounting, Auditing, and Finance* 3 (Fall): 307–39.

Mednick, R. 1987. Accountants' liability: Coping with the stampede to the courtroom. *Journal of Accountancy* 164 (September): 118–22.

Melumad, N., and L. Thoman. 1990. On auditors and the courts in an adverse selection setting. *Journal of Accounting Research* 28 (Spring): 77–120.

Minow, N. 1984. Accountants' legal liability and the litigation explosion. *Journal of Accountancy* 158 (September): 70–86.

Pae, P. 1990. Laventhol says it plans to file for chapter 11. *The Wall Street Journal* (November 20): A1.

Posner, R. 1975a. A theory of negligence. *Journal of Legal Studies.* (1972) Reprinted in *The Economics of Legal Relationships: Readings in the Theory of Property Rights*, edited by H. Manne. Boston: Little Brown, 213–40.

———. 1975b. Strict liability: A comment. *Journal of Legal Studies.* (1973) Reprinted in *The Economics of Legal Relationships: Readings in the Theory of Property Rights,* edited by H. Manne. Boston: Little Brown, 241–57.

Prosser, W., and W. Keeton. 1984. *The Law of Torts.* St. Paul, MN: West Publishing.

Shavell, S. 1980. Strict liability versus negligence. *The Journal of Legal Studies* 9 (January): 1–25.

Smith, V., J. Schatzberg, and W. Waller. 1987. Experimental economics and auditing. *Auditing: A Journal of Practice & Theory* 7 (Fall): 71–93.

Wallin, D. 1992. Legal recourse and the demand for auditing. *The Accounting Review* 67 (January): 121-47.

Washington Update. 1989. *Journal of Accountancy* 169 (August): 109.

Watts, R., and J. Zimmerman. 1983. Agency problems, auditing and the theory of the firm: Some evidence. *Journal of Law and Economics* 26 (October): 613–34.

THE ACCOUNTING REVIEW
Vol. 67, No. 1
January 1992
pp. 121–147

Legal Recourse and the Demand for Auditing

David E. Wallin
The Ohio State University

SYNOPSIS AND INTRODUCTION: Accounting information plays an important role in a decentralized economy. It is a primary way for managers to make assertions about the past performance, current condition, and future prospects of their firms. The auditing of these accounting disclosures is purported to provide value to the economy for several reasons (Baiman 1979; Baiman et al. 1987; Blazenko and Scott 1986; Evans 1980; Scott 1984). Without accurate firm-specific information, investors may align their portfolios in a less-than-optimal fashion, resulting in an inefficient allocation of resources in the society. The public-good feature of auditing may prevent the less efficient private search for firm-specific information. Additionally, auditing is argued to assuage the divergent preferences of managers and investors.

Included in these divergent preferences is managerial motivation to manipulate the accounting information to hide perquisite consumption or imply appropriate production, investment, or financing decisions. Investors anticipate these conflicting goals and price protect themselves, forcing managers to bear the residual loss (Jensen and Meckling 1976). Because of managers' motivation to avoid that loss, auditing can be purchased to provide credibility to managerial disclosures necessary to distinguish among firms of differing quality. With auditing in place, firm managers can reap the benefits of those actions expected to improve the firm's prospects.

Research into managerial incentives to disclose and the effects of disclosures on other managerial actions is at an early stage. The analytical assertions of full disclosure when fraudulent disclosures are not possible are well known (Dye 1985; Grossman 1981; Grossman and Hart 1980;

This article extends a dissertation completed at the University of Arizona under the direction of William S. Waller (chairman), R. Mark Isaac, and Jeffrey Schatzberg. Earlier versions of this research have benefitted from presentations at University of Arizona, Case Western Reserve University, University of Florida, The Ohio State University, University of Pittsburgh, Washington University, and the meetings of the Economic Science Association. The comments of Harry Evans, Ron King, Marla Myers, two anonymous reviewers, and Ira Solomon, the associate editor, are also acknowledged. The funding of the Research Opportunities in Auditing program of KPMG Peat Marwick is greatly appreciated.

Submitted April 1990.
Accepted July 1991.

Milgrom 1981; Milgrom and Roberts 1986), and the body of experimental research is growing (Forsythe et al. 1989; King and Wallin 1990a, 1991). Only recently has either analytical or experimental research been conducted in environments permitting fraudulent disclosure (Dopuch and King 1991, 1992; Dopuch et al. 1989; Kachelmeier 1991; King and Wallin 1990a, 1990b; Wallin 1990).[1] In particular, progress on the formulation of models of the demand for auditing has only recently been made. These models are argued to be at best embryonic (Baiman 1979; Scott 1984) and have typically ignored the possible effects of other mechanisms that may solve the problems that auditing addresses. An important aspect of this article is an investigation of the effect of legal recourse, which allows investors to sue when disclosures are believed to have been fraudulent. The analysis presented here concludes that the threat of lawsuit will cause less frequent fraudulent reports. Importantly, legal recourse also allows the manager to benefit from costly effort, such that the level of effort expended will be that which maximizes societal benefit.

The central purpose of this study is to investigate the demand for auditing in environments both with and without legal recourse. Formal models are developed and tested that investigate the extent to which managers and investors can achieve a cooperative solution when no mechanism exists to aid that cooperation. The analysis and tests are performed in an environment in which the time horizon is not known. This extends the work of Dopuch et al. (1989) by constructing a base-line environment where the demand for auditing is not derived from the backward induction of a single-period equilibrium.

Thirty-two experimental markets were conducted. The results show a demand for auditing, regardless of whether legal recourse was present. The availability of either auditing or legal recourse induced a higher level of managerial effort, the highest occurring when both options were available. Both auditing and legal recourse reduced a tendency for investor overbidding, but only legal recourse reduced the proportion of fraudulent disclosures.

Key Words: *Auditing, Disclosures, Experimental economics, Legal recourse.*

Data Availability: *Data and instructions are available from the author.*

THE article proceeds as follows. The next section develops the theoretical assertions regarding the individual and interactive effects of auditing and legal recourse. Next, the experimental methods are detailed. This is followed by the development of the hypotheses. The results are then presented and the hypotheses are addressed. Finally, the study is summarized.

---

[1] In two related studies, DeJong et al. (1985a, 1985b) concentrated on contracted effort levels, not disclosed outcomes.

I. Theory

There exists a single firm operated by a risk-neutral, effort-averse manager. Initially, assume a single-period world. At the beginning of the period, the manager selects an unobservable effort level. The manager then gains private, imperfect information about the period's earnings and may make a disclosure about that superior information. The right to the current period's earnings is then sold to one of many risk-neutral investors. The manager receives compensation based on the market price of the firm's one share (p^*). For convenience, assume the manager receives p^*. The purchasing investor receives the firm's earnings as a liquidating dividend, once realized. Contracts cannot be formed on outcome.[2]

The firm is a pure security having two possible quality levels, low (q_0) and high (q_1). A low- (high-) quality firm has a q_0 (q_1) probability of high earnings and a $1-q_0$ ($1-q_1$) probability of low earnings. High earnings (s_1) are scaled to one monetary unit and low earnings (s_0) are zero. If the investors know the firm quality to be low (high), they will value the firm at q_0 (q_1).

The manager can affect the firm quality by the selection of either low (e_0) or high (e_1) effort. Low effort is costless; high effort costs ϵ. If the manager selects e_0, there is a π_0 probability of q_1 and a $1-\pi_0$ probability of q_0. The selection of e_1 leads to a π_1 probability of q_1 and a $1-\pi_1$ probability of q_0. The analysis continues with the nontrivial case where $[(1-\pi_1)q_0+\pi_1 q_1]-[(1-\pi_0)q_0+\pi_0 q_1]>\epsilon$.

After selecting the effort level, the manager costlessly observes and may make a disclosure about firm quality (q). Managers may assert that the firm is of high quality (by selecting the disclosure $d=q_1$), assert that it is of low quality ($d=q_0$), or make no assertion ($d=\emptyset$). There is no antifraud rule; the manager may lie. The investors then value the firm at their shared reservation price, p^*, the market price becomes p^*, and the manager receives that amount. After the transaction takes place, all parties observe the actual earning (s). Figure 1 shows the extensive form of the basic game and table 1 summarizes the parameters. The $d=\emptyset$ path has been removed from the extensive-form game for simplicity. The presence or absence of this option will not have an effect on any predicted equilibria.

No Auditing/No Legal Recourse

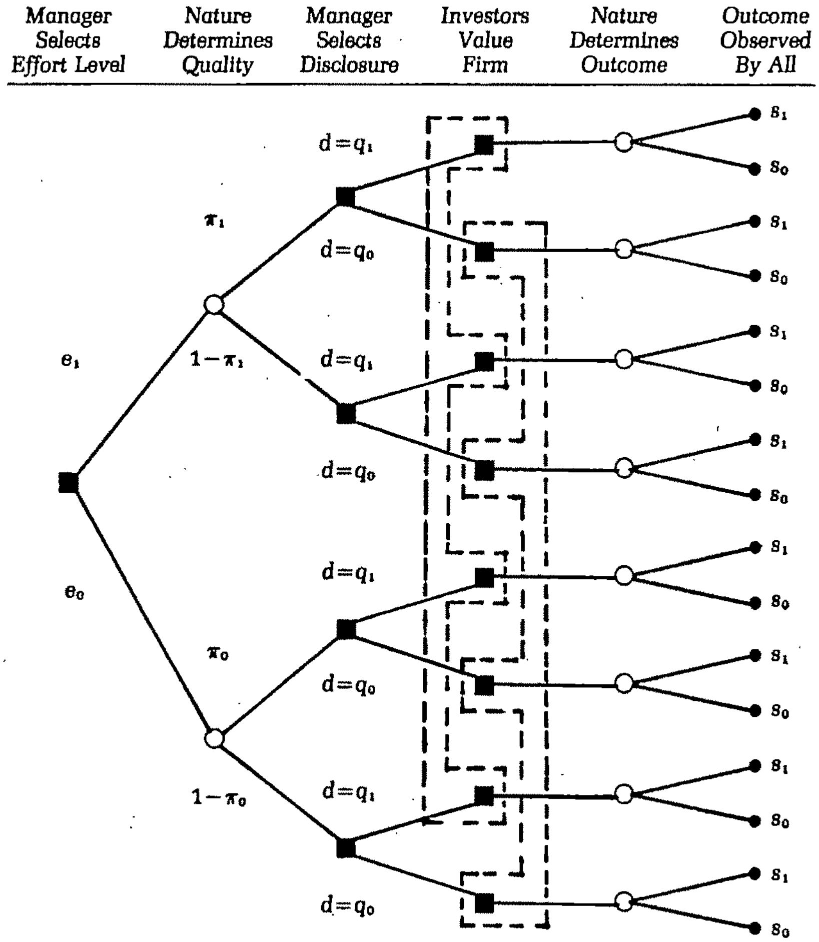
The manager has two decisions to make (effort and disclosure); the investors only determine p^*. Managerial actions depend on the manager's ability to affect p^*. At the pricing node, the only information on managerial action is the disclosure. Since the investors can anticipate that the manager would select the disclosure that produces the largest p^* regardless of q, the disclosure must be uninformative and is ignored.[3] Be-

---

[2] This is consistent with the model of Dopuch et al. (1989) and Dopuch and King (1991, 1992). Kachelmeier (1991) did not allow observation of outcome, thus precluding such contracts. Such contracts could be Pareto-optimal if costs of contracting are sufficiently low. However, in the experiments, this would require complicated contracts, as Pareto-optimality would require the investor with the highest redemption value (defined later) to purchase the asset. These values change each market day and would lead to a quite complex environment. This article's model may be viewed as developing solutions for subgames in which contracting on outcome is prohibitively expensive. The addition of the option of forming such contracts would be an important extension of this work.

[3] King and Wallin (1990a) model this environment and show that the full-disclosure result (Dye 1985; Grossman 1981; Grossman and Hart 1980; Milgrom 1981; Milgrom and Roberts 1986) cannot obtain when there is no antifraud rule. Since the disclosure will be ignored, the manager can select any disclosure policy. However, in

Figure 1
Extensive Form of the Basic Environment



cause the investors cannot observe the manager's effort (or determine it from the disclosure), the manager has a dominant strategy to shirk (i.e., select e_0). Simply put, there is no publicly viewable action that a nonshirking manager can take that a shirking manager cannot copy at no cost. The investors anticipate this and price the firm by assuming e_0 was selected (i.e., $p^*=(1-\pi_0)q_0+\pi_0 q_1$).

Auditing Available/No Legal Recourse

This environment is modeled differently from the agency formulations of the demand for auditing (Baiman 1979; Baiman et al. 1987; Evans 1980). In the analysis devel-

---

conformance with the experimental results of that model (King and Wallin 1990a), it will be assumed that investors believe a manager's report of "bad news." Thus, the manager will report $d=q_1$ whenever disclosures are uninformative.

Table 1
Parameter Definitions

e	Managerial effort, where e_0 is low effort and e_1 is high.
ϵ	Cost of high effort; low effort is costless.
π	Probability that firm is of high quality (q_1). There is a $1-\pi$ probability of having a low-quality firm (q_0). Selecting e_0 always produces π_0; selecting e_1 always produces π_1.
q	Actual firm quality. The firm has q probability of a high outcome (s_1) and a $1-q$ probability of a low outcome (s_0). Firm quality may be low (q_0) or high (q_1).
s	Outcome (earnings) of the firm. The firm may have low (s_0) or high (s_1) earnings.
ν	Verification (auditing) system, where ν_0 designates no purchase and ν_1 designates a purchase.
$d=$	A disclosure (assertion) about firm quality. Three disclosures are possible: $d=q_0$, $d=q_1$, and $d=\emptyset$ (no disclosure).
x	Total amount of legal costs, borne by the loser of the lawsuit.
θ	The probability that the manager will disclose $d=q_1$ when $q=q_0$.
λ	The probability that the investors will bring suit when the manager discloses $d=q_1$ and s_0 obtains.
p^*	The investors' shared reservation value, the market price for the firm's one share, and the managers compensation.

Constraints/ $0<\pi_0<\pi_1<1$, $0<q_0<q_1<1$, $s_0=0$, $s_1=1$
Assumptions: $0\leq\lambda\leq 1$, $0\leq\theta\leq 1$, $x>0$, $(\pi_1-\pi_0)(q_1-q_0)>\epsilon+\alpha$,
Private Information: e, π, q (unless through auditing or legal recourse), and
Public Information: ϵ, s, $d=$, ν, x, p^*, q (in some cases where auditing is purchased or legal action is taken).

oped here, the investors have zero expected profits, leaving the manager to absorb all benefits and costs of auditing and effort (Blazenko and Scott 1986; Jensen and Meckling 1976). The manager can purchase auditing before making the effort selection at a cost of α. The audit report (if released) either agrees or disagrees with the manager's disclosure and is error-free (unlike the strategic, human auditors in Dopuch and King 1991, 1992; Kachelmeier 1991).[4] The investors know that auditing was purchased but receive the audit report only if the manager releases it.[5]

Essentially, the purchase of auditing provides an antifraud rule that can produce an equilibrium with high effort. Although the manager may still lie, it will be to no avail. Consider the subgame in which auditing is purchased. If the manager issues the audit report, the investors will know the firm quality with certainty. A manager observing q_1 will disclose it as $d=q_1$ with an "agree" report, obtaining maximum firm price (q_1) with certainty.[6] If the manager does not issue the audit report, the investors correctly reason that q_0 obtained. Because investors know the actual quality when auditing is purchased, the expected compensation to a manager selecting e_0 (e_1) is $(1-\pi_0)q_0+\pi_0 q_1$ $((1-\pi_1)q_0+\pi_1 q_1)$. The net-of-effort increase in expected compensation by selecting e_1 is $(\pi_1-\pi_0)(q_1-q_0)-\epsilon$ (given previous assumptions, this is greater than zero). Thus, the manager will not shirk when auditing has been purchased and the investors will know firm value with certainty.

---

[4] Both auditing and the legal system work without error. This seems a reasonable starting point for early exploration. The availability of imperfect auditing and/or legal recourse (whether mechanistic or strategic) is an interesting future extension.

[5] There is no effect on the equilibria if the audit purchase is or is not public information or if the report can or cannot be withheld by the manager. The manager could be allowed to purchase auditing before or after the effort decision and before or after q is observed without a qualitative change in the equilibrium predictions. The only quantitative effect would come from a delay in purchase until after q is observed, since the manager would save α when q_0 obtains.

If auditing has not been purchased, the prediction of shirking from the previous section holds. If the manager does (does not) purchase auditing, net-of-effort managerial compensation is $(1-\pi_1)q_0+\pi_1 q_1-\epsilon$ $((1-\pi_0)q_0+\pi_0 q_1)$. The auditing considered here is "reasonably priced" (i.e., $(\pi_1-\pi_0)(q_1-q_0)-\epsilon>\alpha$) and will be purchased in equilibrium.

No Auditing/Legal Recourse Available

King and Wallin (1990b) and Wallin (1990) considered the effects of legal recourse that correctly concludes guilt. The manager always discloses $d=q_1$ when $q=q_1$ and asserts $d=q_1$ when $q=q_0$ with probability Θ. The purchasing investor would consider bringing suit only when the manager issues $d=q_1$ and s_0 obtains, doing so with a probability of λ. If the investor sues and the manager has lied (i.e., has issued $d=q_1$ when $q=q_0$), the investor is awarded damages from the manager equal to p^*-q_0.[7] The total legal costs of both the manager and investor, $\varkappa$, are borne by the loser of the lawsuit (i.e., the British method of allocating legal costs).[8]

Assume legal costs are not sufficiently high to preclude legal action. There exist mixed-strategy equilibria on reporting and suing strategies. King and Wallin (1990b) and Wallin (1990) show that expected managerial lying is:

$$\Theta=\frac{\varkappa\pi(1-q_1)}{(1-\pi)[(q_1-q_0)(1-q_0)-\varkappa(1-q_1)]}, \tag{1}$$

and that the proportion of suing will be in the range:

$$\frac{(q_1-q_0)(1-q_0)-\varkappa(1-q_1)}{(q_1-q_0)(1-q_0)(1-q_0+\varkappa)}$$
$$\geq\lambda\geq\frac{(q_1-q_0)[(q_1-q_0)(1-q_0)-\varkappa(1-q_1)]}{(1-q_0)[(q_1-q_0)\{\varkappa(q_1-q_0)+q_1\}+(1-q_0)q_0^2]}. \tag{2}$$

---

[6] For convenience, it is assumed that the manager will not issue $d=q_0$ with a "disagree" audit report, although it effectively serves the same purpose. This assumption has no effect on the equilibrium predicted here and is not necessary in the environment with legal recourse. No subject paired $d=q_0$ with a "disagree" audit report in the late market days described later.

[7] Except for a reimbursement for legal fees, current U.S. securities laws allow a recovery of actual (out-of-pocket) damages only. The damages are computed as the additional amount paid by the investor because of the fraud. Rather than calculate damages equal to p^* (i.e., a rescission amount or refund of the purchase price), the damages are p^*-q (a refund of the "excessive" purchase price) since the investor would have paid q if the disclosure was constrained to be the truth. Thus, no recovery is possible when the disclosure $d=q_0$ is selected, since $p^*=q_0$ in that case. Since auditing is perfect, there would never be action against the auditors (in contrast to Dopuch and King 1991, 1992).

[8] The qualitative results are unaffected by the use of the American system in which litigants pay their own legal fees regardless of outcome. Of the two, the relevant incentives in the British system more closely resemble those in current U.S. securities law.

At equilibrium, the manager and investors will know Θ, the range of λ, and π. The threat of a lawsuit both reduces the frequency of lying (since $\Theta<1$) and allows the manager to reap the entire benefit of higher effort (net of expected legal costs). Therefore, the manager will exert high effort.

Auditing Available/Legal Recourse Available

The solution when both legal recourse and auditing are available depends on the relative levels of their costs. Although it is reasonable to assume that legal costs include a full audit (i.e., $\varkappa>\alpha$), conclusions in this environment depend on the relative magnitude of the parameters. A discussion of specific predictions will be left until the parameters are assigned values later. However, the general observation can be made that, unless $\varkappa$ is much larger than α, the manager is better off with no audit purchases, trading potential losses from litigation for the certain cost of auditing.

Multiperiod Environments

The previous theory was developed with an assumption of a single-period world.[9] When there are multiple periods, there can be alternate equilibria. If it is known which period will be the last, the equilibria predicted here would hold in that last period. Backward induction would allow the extension of the single-period equilibria to multiperiod environments with a known endpoint. In a game with an unknown horizon (i.e., where, for any single period, the probability of moving on to the next period is always greater than zero), backward induction cannot be used. Although single-period equilibria carry over to the unknown-horizon world, new equilibria can exist. An important feature of this study is the use of a base-line environment that does not use the backward induction of a single period to obtain the demand for auditing.

The basic environment analyzed here is essentially a prisoner's dilemma. Both the manager and investors would be better off (or no worse off) with an equilibrium involving high effort and a commensurate market price. Yet, the dominant single-period strategy for the manager is to shirk. With an unknown horizon, this game can have a Pareto-efficient, cooperative Nash equilibrium or any of the infinite partially cooperative, mixed-strategy Nash equilibria (Klein and Leffler 1981; Ordeshook 1986; Townsend 1982).[10] Given the many possible equilibria, game theorists typically make assumptions about the heuristics used by players in such a game to derive an equilibrium (Ordeshook 1986).

A commonly used heuristic is the tit-for-tat strategy (Axelrod 1984). A player using this strategy cooperates in the first period. In subsequent periods, the tit-for-tat player matches the other player's choice from the preceding period. Although the manager in this environment is immediately aware of the investors' action (i.e., the market price), the investors do not know e. By observing the stream of s over multiple periods, the investors can make inferences about the frequency of e_1 selection. Because investors cannot determine immediately when a defection (e_0) has taken place, this environment may differ from the typical unknown-horizon prisoner's dilemma. Bendor (1986, 1987) considered a similar game in which players' actions map onto observed outcomes through a stochastic process (i.e., there is "noise").

---

[9] This multiperiod environment is developed much like a "snapshot" of an overlapping generations model (Dye 1990).

[10] The equilibria can be affected by the discount factor used by the players. As the discount factor approaches infinity, the players discount the future so much that the game essentially becomes a one-period game. This analysis will assume the discount factor is reasonably small compared to the magnitude of other parameters. For notational convenience, the future will not be discounted.

Bendor (1986, 1987) found that two tit-for-tat players paired in a prisoner's dilemma with noise will not have a cooperative solution each period. Players "punish" each other for perceived defections that are simply bad states of nature. Thus, the prisoner's dilemma with tit-for-tat strategies and noise stabilizes with partial cooperation. Although this analysis may provide insight for the development of a solution to the basic environment, the implementation of a tit-for-tat strategy in a market may be difficult.[11] A possible aid to the coordination problems is the use of truthful disclosures. The subgame-perfect strategy of uninformative disclosures would hold in a single-period world, but the existence of an unknown time horizon does not preclude an equilibrium with truthful disclosure, particularly if truthful disclosure is used to sustain a high-effort equilibrium. A manager who consistently selects e_1 may find it necessary to disclose observances of q_0 truthfully to convince investors that it was chance, not defection, that increased the likelihood of s_0.

Roth and Murnighan (1978) and Murnighan and Roth (1983) developed and tested predictions about cooperative behavior in a repeated prisoner's dilemma with an unknown horizon. In a game without noise, subjects (possibly lacking salient rewards) selected the cooperative strategy less than 40 percent of the time, despite a high likelihood (90 percent) of continuation of the game. For this reason, and in light of previous results without an unknown horizon (Dopuch et al. 1989; Kachelmeier 1991), as well as the addition of noise, a maintained hypothesis for the remainder of this development will be that low levels of cooperation will be found when no auditing or legal recourse is available, permitting auditing or legal recourse to have value.

II. Method

The markets in two cells share most features with those of Dopuch et al. (1989), who investigated the manager's and investors' demand for auditing in an environment without legal recourse using a known time horizon. The markets in the basic and auditing-only environments (cells A and B) differ from those of Dopuch et al. (1989) primarily in the time-horizon assumption. This provides a proven mechanism and set of parameters that also generates evidence on time-horizon effects. Each market consisted of one seller (manager) who offered an asset—lottery "ticket" (the firm)—for sale to three buyers (investors) in an environment with an unknown time horizon.

Experimental Design

In the experimental design presented below, the presence or absence of auditing was contrasted with the presence or absence of legal recourse for investors. Eight markets were conducted in each cell and the subjects were paid student volunteers.

	Auditing Unavailable	*Auditing Available*
No Legal Recourse	A	B
Legal Recourse	C	D

---

[11] The competitive nature of the environment requires that p* reflect the expected managerial effort for this period, not the effort (or outcome) from the previous period. A tit-for-tat investor strategy would require that the investors effectively tacitly collude. Tacit (even explicit) collusion is rarely maintained in experimental markets (e.g., see Isaac and Walker 1986).

Table 2
Sequence of Markets

First Series			Second Series		
First Market	*Second Market*	*Seller*	*First Market*	*Second Market*	*Seller*
A1	B1	Same	A5 A6	A7 A8	New (4 different sellers)
A2	C1	Same			
B2	A3	Same	B5 B6	B7 B8	New (4 different sellers)
B3	D1	Same			
C2	A4	Same	C5 C6	C7 C8	New (4 different sellers)
C3	D2	Same			
D3	B4	Same	D5 D6	D7 D8	New (4 different sellers)
D4	C4	Same			

For the first series of markets (those numbered 1–4), subjects (University of Arizona undergraduate business majors) were recruited in groups of four, and upon arrival, one subject in each group was randomly assigned the role of seller. After completing a market in one cell, the subjects were assigned the same roles in a second market from a different cell. Table 2 shows the first and second markets for these groups. The following cell pairings were used: AB, BA, AC, CA, BD, DB, CD, DC. No AD, DA, BC, or CB pairings were conducted because a change in both auditing and legal recourse seemed unnecessarily complex.

In the second series of markets, subjects (undergraduate business majors at The Ohio State University) were recruited in groups of eight and, upon arrival, were randomly assigned as seller or buyer in one of two groups. Both groups participated in separate markets from the same cell. After the first market, the group of eight was again randomly assigned to groups and roles under the announced constraint that a subject could be a seller only once and that a seller-turned-buyer from one first-market group could not remain in the same second-market group as a buyer-turned-seller.[12] The two groups then experienced a separate, second market from the same cell as the first market.

[12] The changes in the method of data collection were made for two reasons. First, logistic considerations and computer lab availability dictated that groups of eight would be necessary. The seller could be seated in a separate room in the first series of markets. This could not be done in the second series. However, by conducting two simultaneous markets in the second series, the subjects did not know with whom they were grouped. Second, the initial design was constructed to allow for an observation of the effect of a change in auditing or legal recourse availability for the same seller. Analysis of that data suggested no additional information was present. The 32 markets reported had 24 different subjects acting as sellers.

Market Session

The sequence of the session was: (1) instructions and training for the first market, (2) the first market, (3) instructions (if different cell) and training for the second market, (4) the second market, and (5) the post-experimental phase. No information about the second market was given before the conclusion of the first. The two markets were separate units, except that learning and some relationships established in the first market may have carried over to the second.

Instructions and Training. A computer program displayed the instructions for the market. Following the instructions, a training phase (no cash earned) began that consisted of five experimental market "days" that used the same computerized trading mechanism as the experimental phase. The purpose of this training phase was to allow the subjects to learn about the computer mechanism, not the behavior of others. For this reason, the experimenter gave the subjects the values to enter. Subjects were informed that these values were to illustrate the options available to them and were not to be considered as recommended behavior.

The Market. Each market had an unknown time horizon and lasted at least 25 market days. On day 25 and after, the experimenter rolled a pair of dice in the presence of the subjects. The market continued unless the dice matched.[13] The subjects were informed of the two-hour time limit for the session and that, should a terminating roll not be observed within that time limit, they would be brought back to continue the session. This was never necessary.

Each day the seller had one asset available for sale. The asset could be one of two quality levels. A high (low) quality asset had a 90 (10) percent chance of becoming a winner and a 10 (90) percent chance of becoming a loser. A losing asset was worthless; a winning asset had a value that depended on the buyer's redemption value: $2.80, $2.00, or $1.20. At the beginning of each market day, each buyer was assigned one of the three redemption values, each equally likely. The random assignment was made with replacement, so two or three buyers were often assigned the same redemption value. All subjects were informed of the method of determining value and the amount of the possible redemption values. Multiple redemptions values allowed the buyers to have positive expected profits despite seller effort. However, (as it will be shown later) expected profits to buyers were higher when high effort was exerted.

The asset was sold by using a second-price, sealed-bid institution: the buyer submitting the highest bid received the asset at the amount of the second highest bid (Cox et al. 1982). Although the analytical formulation did not specify an institution, the second-price, sealed-bid auction is consistent with the analytical assumptions. The dominant strategy for a buyer is to bid his reservation value: a risk-neutral buyer would bid 90 (10) percent of the redemption value if the asset was known to be of high (low) quality. A buyer's belief that an asset was of high quality could then be deduced from the ratio of the bid to the redemption value.

The expected value of the highest (second highest) redemption value was $2.53 ($2.00). If the asset was known to be of high quality, the expected price would be $1.80 (90 percent of $2.00) and expected profits of the winning buyer would be $0.48 (90 percent of \$2.53 − \$2.00). The expected price of an asset known to be of low quality would be \$0.20 (10 percent of \$2.00) and expected profits of the winning buyer wold be \$0.05 (10 percent of \$2.53 − \$2.00). Thus, buyers had higher expected profits from high-quality assets.

---

[13] Thus, there was a one-sixth chance of the market ending any day after day 24. This essentially produced a maximum discount factor of 20 percent $[p/(1-p)]$ for the subjects. The parameters were selected with this in mind. The expected number of experimental days was 30 per market.

The seller had two choices regarding effort, implemented as "bins." When the "default bin" (corresponding to low effort) was selected, there was a 20 percent chance the asset would be of high quality and an 80 percent chance of low quality. The "alternate bin" (high effort) had an 80 percent chance of high quality and a 20 percent chance of low. The default bin could be selected without cost; the alternate bin cost the seller \$0.40. Thus, the probability that an asset from the default bin would win was 26 percent ($0.2 \times 0.9 + 0.8 \times 0.1$) and, for the alternate bin, it was 74 percent ($0.8 \times 0.9 + 0.2 \times 0.1$). If no information asymmetry on actual quality existed (or actual effort selection), the increase in expected price would be \$0.96 ((0.74–0.26) × \$2.00). In an equilibrium with high effort and commensurate pricing, the expected, net-of-effort profits to the seller would be \$0.56 (\$0.96 − \$0.40) higher than in the low-effort equilibrium.

The available actions and rules in steps 1–9 (as applicable to the cell under study) were common knowledge and were:

1. Buyers were randomly assigned (with replacement) redemption values.
2. *Cells B and D only*: Seller selected whether to purchase auditing. Buyers were informed of this choice.
3. Seller selected either the default or alternate bin. Actual quality was determined and observed by the seller.
4. Seller selected and buyers observed disclosure.
5. *Cells B and D only*: Seller selected whether to release the audit report if auditing was purchased. Buyers observed the audit report or message that none was forthcoming.
6. Buyers submitted bids to purchase asset.
7. Winning bidder was determined, bids were displayed, the lottery was played, and the outcome (won or lost) of the lottery was displayed.
8. *Cells C and D only*: If the asset was disclosed as high and lost, the buyer who purchased it selected whether to bring suit. If litigation took place, the winner of the lawsuit was determined.
9. Cash balances were updated.

In step 1, redemption values were assigned as previously detailed. In step 2, the seller could purchase auditing (called verification) at a cost of \$0.30 in cells B and D. The auditing was "reasonably priced," as the expected, net-of-effort, net-of-auditing, increased profits to the seller were \$0.26 (\$0.56 − \$0.30). In cells B and D, the buyers were informed whether auditing had been purchased.

In steps 3 and 4, the seller selected between bins. After this selection was made, the computer used the priors from the selected bin to determine randomly the actual quality and the seller observed that quality. Then, the seller selected among three disclosures: "high," "low," or "either" (i.e., no disclosure was made). The seller could select any of the three, despite actual quality (i.e., sellers could lie). The buyers then observed the disclosure.

In step 5, if the seller in a cell B or D market did not purchase auditing, the buyers were reminded of that selection. If the seller in a cell B or D market purchased auditing, the seller determined whether to send an audit report (called a verification signal) with the disclosure. If the audit report was not sent, the buyers were informed that no report was forthcoming. If the seller sent the audit report, it would accurately "agree" or "disagree" with the seller's disclosure.

By step 6, buyers had the seller's disclosure. In markets in cells B and D the buyers also knew (1) auditing was not purchased, (2) auditing was purchased but not reported, or (3) the audit report. The buyers then submitted bids to purchase the asset. If two or more buyers tied for the highest bid, the computer randomly selected the winner from those tied bidders. All parties observed the winning bid, losing bids, and the market price. The computer then used the prior probabilities derived from the actual quality to determine randomly whether the asset was a winner or loser. All parties were informed of whether the asset won or lost.[14]

In step 8, the purchasing buyer in cells C and D had the option to bring suit against a seller when a losing asset had been disclosed as being of high quality. When these conditions occurred, the purchasing buyer was prompted to select whether to bring suit (called investigation) and was informed that: (1) no suit brought about no further change in cash balances, (2) undertaking and winning a lawsuit led to a reimbursement from the seller to buyer and the seller lost $0.50 in legal costs (called investigation costs), and (3) undertaking and losing the lawsuit resulted in a loss of $0.50 to the buyer and no change for the seller. Legal costs were a dead-weight loss.[15] The reimbursement was the lesser of (1) the market price less 10 percent of the buyer's redemption value, or (2) 80 percent of the buyer's redemption value (an "actual damages" amount).[16] All parties were informed of the selection made by the purchasing buyer and the results of the lawsuit.

In step 9, cash balances were updated and the computer program prepared for the next market day (subject to the dice roll). The computer program allowed the results of the previous 14 market days to remain on the display screen.

At the end of each market day, all actions and information sets were common knowledge except the buyers were never informed of which bin (effort) was selected by the seller. Also, although actual quality could be determined by the buyers from the audit report (in cells B and D) or from the result of a lawsuit (in cells C and D), buyers otherwise would not know (for certain) what quality obtained. Redemption values were private information except that the seller and other buyers could determine the redemption value of a buyer winning a lawsuit (in cells C and D only). Cash balances were kept private.

Post-Experimental Phase. Cash accumulated separately for each market. Each subject began each market with $8 in cash. The seller's cash balance was increased by the sales price of each asset and was decreased by bin (effort) costs, auditing costs (when purchased), and losses from lawsuits (both damages and legal costs upon losing). A buyer's cash balance was increased by the redemption value of a winning asset and by actual damages received from a lawsuit. The buyer's cash balance was decreased by the price of assets purchased and legal costs upon losing a lawsuit.[17] After the first market, the cash was "banked." All parties were informed that the seller received a cash payment equal to 25 percent of his ending cash balance and each buyer received a cash payment equal to 75 percent of his ending cash balance. The average compensation for buyers was approximately $14 each and $18 for the seller.

---

[14] Dopuch et al. (1989) found a significant demand for auditing only when there was a delay in observing the outcome (s). Kachelmeier's (1991) investors (owners) did not learn any outcomes until the end of the market.

[15] This would produce a predicted lying rate (Θ) of 14 percent for a seller selecting the alternate bin. The lying rate would drop to 1 percent for a seller selecting the default bin.

[16] The 80 percent amount represents the maximum risk-neutral price for a high-quality asset, 90 percent, reduced by the risk-neutral value of a low-quality asset (10 percent). This is the maximum actual damages amount.

III. Hypotheses

The predicted equilibrium behavior in three of the four cells are extremes. For example, the manager in cell B was predicted to purchase auditing and exert high effort always. Standard statistical tests would require a rejection of a hypothesis of 100 percent high effort with a single contrary observation. This rejection would be expected, as experimental results rarely support such behavioral extremes. The construction of a base-line market (cell A) and the 2×2 design allows a comparison of behaviors as the presence of auditing and legal recourse are manipulated. The central purpose of this study is to compare the effects of auditing and legal recourse on behaviors, not necessarily to test whether a particular extreme equilibrium holds in a selected cell. Thus, where possible, the hypotheses are comparative in nature.[18]

The first series of hypotheses deals with effort selection. From the previous analysis and the experimental design, the following predictions can be made.

H_{1a}: The proportion of high effort is greater when auditing is available than when it is not.

H_{1b}: The proportion of high effort is greater when legal recourse is available than when it is not.

H_{1c}: When both auditing and legal recourse are available, the proportion of high effort is no greater than when auditing alone or legal recourse alone is available.

The second series of hypotheses deal with audit purchases. Although auditing was predicted to be purchased when legal recourse was not present, it is less expensive (with the selected parameters) for the manager to suffer expected legal costs than to purchase auditing consistently when both mechanisms exist. Thus,

H_{2a}: When no legal recourse is available, auditing is purchased when available.

H_{2b}: When auditing is available, the proportion of audit purchases is less when legal recourse is available than when it is not.

The third series of hypotheses relates audit purchases with effort selection. Without legal recourse, auditing can be purchased to move toward the preferred equilibrium.

---

[17] Cash balances were not constrained to be positive. No participant's cash balance fell to a point that precluded full participation in any market day for fear of "bankruptcy."

[18] It should be noted that the market mechanism was designed purposefully to work against most of these hypotheses. By allowing immediate feedback on whether the asset won (i.e., s is known to investors at period's end), this environment allows a high-effort manager to more easily form a reputation for exerting high effort. Dopuch et al. (1989) and Kachelmeier (1991) found a significant demand for auditing only when the observation of s was delayed. Thus, the demand for auditing and the increase in effort from auditing or a legal system should be at its lowest level here compared with an environment where s is delayed.

However, the presence of a legal system allows a higher proportion of high effort to exist without auditing. Thus,

H_{3a}: When no legal recourse is available, the proportion of high effort is greater when auditing is purchased.

H_{3b}: When legal recourse is available, the proportion of high effort is unaffected by the presence of auditing.

The fourth hypothesis deals with the level of fraudulent disclosure. In the previous analysis, the availability or purchase of auditing does not necessarily affect the truthfulness of disclosures. However,

H_4: The proportion of fraudulent disclosure is lower when legal recourse is present.

The final series of hypotheses deals with economic efficiency. Economic efficiency is a measure often considered in experimental markets and details the amount of resources generated by trade. In this setting, the maximum economic efficiency comes when the manager exerts high effort and the investor with the highest redemption value purchases the asset. The hypotheses address gross efficiency and the costs of auditing or the legal system are not considered. Net efficiency, after such costs reduce gross efficiency, is considered in the results section. The analysis predicts:

H_{5a}: Economic efficiency is higher when auditing is available than when it is not.

H_{5b}: Economic efficiency is higher when legal recourse is available than when it is not.

H_{5c}: When both auditing and legal recourse are available, economic efficiency is no higher than when auditing alone or legal recourse alone is available.

IV. Results

Figure 2 shows the proportion of high effort within each cell changed over time. To allow for learning, the market days were divided into two groups: early (days 1–15) and late (days 16 and later).

Tables 3–6 present the results for each cell, including a segregation into early and late market days. However, the discussion and all statistical tests deal with late market days.[19] The tables present the proportion of high effort (alternate bin selection) and the proportion of truthful reports by the seller. A truthful report occurred when the seller observed a low-quality asset and did not report it as high. A bid ratio was computed for each buyer as that buyer's bid divided by that buyer's redemption value for that day. The median of the bid ratio was then calculated for each day. The mean bid ratio was the mean of this measure across the relevant days. The mean bid ratio is used later to compare the buyers' ability to match bids to underlying quality. The results for cells with legal recourse (C and D) also contain a listing of the frequency of lawsuits. For example, the buyers in the early days of market C4 brought suit in three of the five opportunities available (disclosure of high quality for an asset that lost) or 60 percent of the time. Table 7 summarizes effort and truthfulness in all markets.

---

[19] The separation at day 15 is, of course, somewhat arbitrary. The graph does suggest that, at least for effort selection, a distinct separation occurs after the first ten days. The statistical tests are generally unaffected by differing definitions of the separation day.

Figure 2
Effort Selection over Time

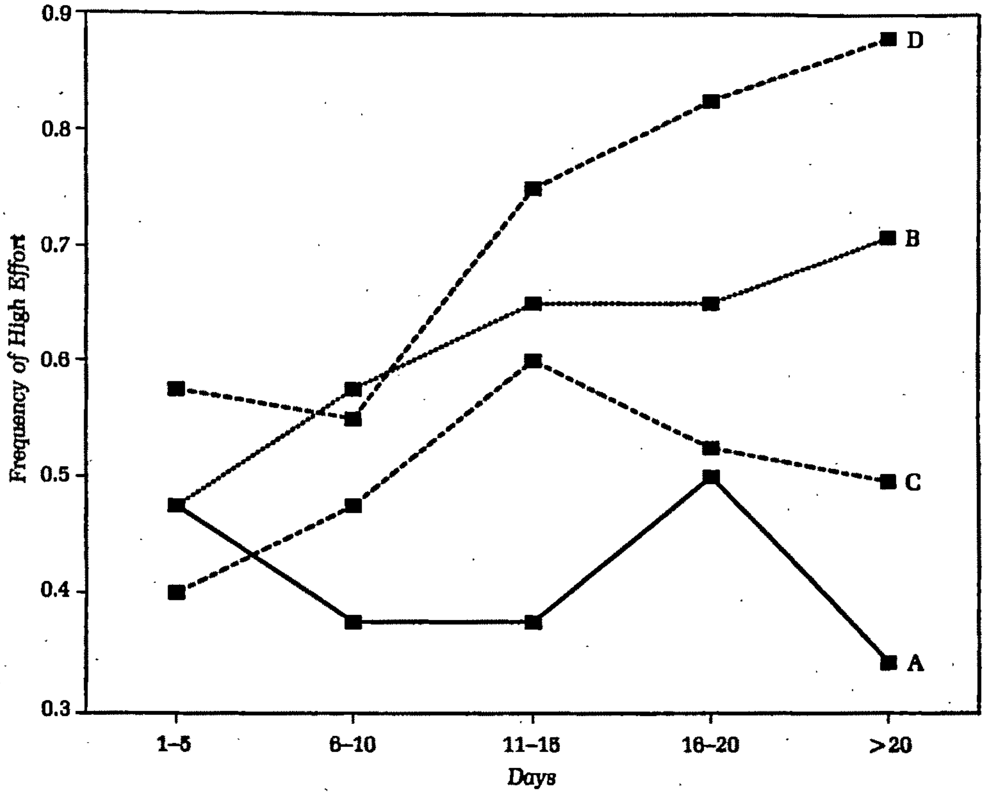


Table 3
Cell A Results
No Auditing/No Legal Recourse

	N		*Percentage of High Effort*		*Frequency of Truthful Disclosure*		*Percentage of Wins*		*Mean Bid Ratio*	
Market	*Early*	*Late*	*Early*	*Late*	*Early*	*Late*	*Early*	*Late*	*Early*	*Late*
A1	15	15	40.0	33.3	6/9*	4/9	33.3	53.3	52.4	65.4
A2	15	12	86.7	33.3	2/4	2/7	80.0	41.7	59.8	78.3
A3	15	13	40.0	61.5	6/8	0/5	53.3	61.5	63.9	81.4
A4	15	24	66.7	91.7	3/3	7/9	73.3	62.5	92.9	78.0
A5	15	13	26.7	38.5	6/9	7/8	33.3	30.8	28.1	28.4
A6	15	10	33.3	20.0	2/8	5/7	46.7	30.0	44.2	40.8
A7	15	18	26.7	22.2	5/8	10/14	46.7	27.8	36.4	49.0
A8	15	11	6.7	18.2	10/10	8/9	46.7	36.4	46.2	41.0
Average			40.8	39.8	68.2%	58.8%	51.7	43.0	53.0	57.8

* That is, the seller had nine occurrences of low quality and disclosed six of them as low or either.

Table 4
Cell B Results
Auditing Available/No Legal Recourse

Market (Audit Choice)	N		*Percentage of Audit Purchases*		*Percentage of High Effort*		*Frequency of Truthful Disclosure*		*Percentage of Wins*		*Mean Bid Ratio*	
	Early	*Late*	*Early*	*Late*	*Early*	*Late*	*Early*	*Late*	*Early*	*Late*	*Early*	*Late*
B1 (No)	10	7	33.3	30.0	50.0	28.6	2/6	2/5	50.0	28.6	61.6	59.9
(Yes)	5	3			100.0	100.0	—	—	100.0	100.0	86.3	99.4
B2 (No)	2	5	86.7	66.7	50.0	40.0	1/1	0/4	50.0	20.0	62.7	75.4
(Yes)	13	10			23.1	50.0	3/8	1/3	53.8	70.0	65.6	83.4
B3 (No)	11	25	26.7	7.4	90.9	100.0	1/4	3/5	72.7	76.0	70.7	65.0
(Yes)	4	2			50.0	100.0	4/4	—	0.0	100.0	33.7	53.6
B4 (No)	13	7	13.3	50.0	61.5	85.7	4/4	1/1	61.5	71.4	31.3	42.0
(Yes)	2	7			100.0	85.7	0/1	—	50.0	100.0	57.9	98.5
B5 (No)	7	4	53.3	69.2	57.1	100.0	4/4	1/1	28.6	75.0	31.0	33.1
(Yes)	8	9			87.5	100.0	—	2/2	87.5	55.6	77.9	74.0
B6 (No)	8	6	46.7	53.8	37.5	16.7	3/5	0/4	37.5	50.0	53.4	77.4
(Yes)	7	7			71.4	100.0	2/3	1/1	57.1	71.4	65.6	85.6
B7 (No)	9	9	40.0	50.0	44.4	44.4	0/3	1/3	44.4	44.4	52.9	64.9
(Yes)	6	9			66.7	55.6	2/3	3/4	50.0	55.6	73.8	60.0
B8 (No)	7	4	53.3	69.2	0.0	0.0	2/5	0/4	14.3	0.0	60.3	28.0
(Yes)	8	9			62.5	77.8	2/4	0/2	37.5	77.8	65.3	77.1
Average (No)			44.2	49.5	48.9	51.9	57.3%	41.7%	44.9	45.7	53.0	55.7
Average (Yes)					70.1	83.6	53.5	61.7	54.5	78.8	65.8	79.0
Average (All)					59.5	67.8	55.7	49.4	49.7	62.2	59.4	67.3

Tests of Hypotheses

In all statistical tests, a single observation was derived from the late market days within each market or subdivision of each market (i.e., separated into audit and non-audit days). Since two locations were used and markets were paired in various orders, tests were performed that found no significant threat from pooling the data.[20] Also, subjects showed no significant evidence of a deviation from risk neutrality.[21]

Table 5
Cell C Results
No Auditing/Legal Recourse Available

	N		*Percentage of High Effort*		*Frequency of Truthful Disclosure*		*Percentage of Wins*		*Mean Bid Ratio*	
Market	*Early*	*Late*	*Early*	*Late*	*Early*	*Late*	*Early*	*Late*	*Early*	*Late*
C1	15	13	86.7	92.3	6/7	4/5	53.3	61.5	77.2	79.5
C2	15	13	40.0	69.2	7/9	2/3	40.0	69.2	55.6	85.4
C3	15	10	33.3	50.0	8/10	6/7	33.3	40.0	65.7	78.6
C4	15	13	33.3	61.5	3/5	3/7	46.7	69.2	65.4	53.0
C5	15	14	60.0	57.1	5/6	9/9	60.0	57.1	62.3	62.8
C6	15	10	33.3	0.0	8/9	4/7	40.0	30.0	69.1	61.1
C7	15	14	46.7	35.7	6/8	6/6	40.0	57.1	69.1	57.7
C8	15	12	60.0	41.7	6/8	7/9	46.7	41.7	59.6	59.2
Average			49.2	51.0	78.2%	79.0%	45.0	53.2	65.5	67.2

	Frequency of Legal Action			
Market	*Early*		*Late*	
C1	100.0%	(2 of 2)	100.0%	(1 of 1)
C2	50.0	(1 of 2)	100.0	(2 of 2)
C3	100.0	(2 of 2)	—	—
C4	60.0	(3 of 5)	66.7	(2 of 3)
C5	100.0	(1 of 1)	—	—
C6	100.0	(2 of 2)	100.0	(3 of 3)
C7	100.0	(3 of 3)	—	—
C8	100.0	(2 of 2)	—	—
Average	88.8%		91.7%	

---

[20] Three tests used a MANOVA to determine the appropriateness of the pooling of the data. A MANOVA was used to provide the best chance of detection of possible effects beyond that observed in univariate tests. These tests used effort, truthfulness, and percentage of over- or underbidding as dependent variables and the presence/absence of auditing, presence/absence of legal recourse, and the interaction of the two as factors. Overall, both auditing ($p<0.003$) and legal recourse ($p<0.007$) had a significant effect on the three dependent variables. The interaction of auditing and legal recourse was not significant ($p>0.7$). A MANOVA was constructed with an additional variable indicating whether the market was the first or second market in the session. This variable (and interactions involving this variable) never had a significant effect on the dependent variables (all $p>0.4$). A similar test of a difference between the first four and second four markets in each cell showed a similar, low threat (all $p>0.2$) from pooling the data.

[21] The analytical development modeled risk-neutral parties. The experimental design intentionally incorporated divergent parameters to protect the qualitative predictions from the threat of non-risk-neutral behavior. For example, the *ex ante* probability that high (low) effort leads to a winning lottery is 74 percent (26 percent). Thus, risk preferences would have to diverge greatly from risk neutrality for the qualitative predictions to change. Although non-risk-neutral preferences could explain why the equilibrium behaviors were not at the extremes predicted by the analytical developments, the subjects' behavior appeared consistent with near risk-neutral behavior. On market days in cells B and D when the presence of an agree or disagree verification report made the quality of the asset known to all, the investors' bids were not significantly different from the risk-neutral prediction ($p>0.2$).

Table 6
Cell D Results
Auditing Available/Legal Recourse Available

Market (Audit Choice)	*N*		*Percentage of Audit Purchases*		*Percentage of High Effort*		*Frequency of Truthful Disclosure*		*Percentage of Wins*		*Mean Bid Ratio*	
	Early	*Late*	*Early*	*Late*	*Early*	*Late*	*Early*	*Late*	*Early*	*Late*	*Early*	*Late*
D1 (No)	11	12	26.7	0.0	100.0	100.0	1/2	2/3	72.7	66.7	70.0	67.8
(Yes)	4	—			75.0	—	1/1	—	75.0	—	53.7	—
D2 (No)	12	8	20.0	38.5	50.0	50.0	6/8	4/5	41.7	25.0	70.9	76.2
(Yes)	3	5			100.0	100.0	2/2	—	33.3	80.0	75.8	92.6
D3 (No)	9	4	40.0	60.0	44.4	100.0	3/4	—	16.7	83.3	23.8	31.8
(Yes)	6	6			50.0	100.0	5/5	1/1	100.0	28.6	26.2	84.6
D4 (No)	7	10	53.3	63.0	14.3	40.0	5/7	3/3	28.6	50.0	59.3	50.1
(Yes)	8	17			75.0	82.4	2/2	2/3	75.0	82.3	84.5	81.3
D5 (No)	14	15	6.7	0.0	50.0	93.3	5/5	5/5	71.4	66.7	62.3	67.8
(Yes)	1	—			100.0	—	1/1	—	0.0	—	30.0	—
D6 (No)	10	4	33.3	60.0	50.0	75.0	3/5	1/1	50.0	75.0	51.0	73.0
(Yes)	5	6			40.0	83.3	3/3	1/1	40.0	100.0	55.4	90.0
D7 (No)	15	15	0.0	16.7	93.3	86.7	4/4	8/8	73.3	53.3	69.6	86.6
(Yes)	—	3			—	100.0	—	—	—	100.0	—	99.7
D8 (No)	13	12	13.3	14.3	53.8	91.7	5/5	2/2	61.5	91.7	62.5	82.9
(Yes)	2	2			100.0	100.0	—	1/1	100.0	50.0	94.6	50.4
Average (No)			24.2	31.5	57.0	79.6	78.9%	92.4%	56.9	66.0	58.7	67.0
Average (Yes)					77.1	94.3	100.0	91.7	48.6	82.6	60.0	83.1
Average (All)					66.4	85.9	88.0	92.1	53.0	73.1	59.3	73.9

(continued)

To test the first series of hypotheses, an ANOVA was constructed[22] that had the proportion of high effort in each market as the dependent variable (giving 32 observations). The factors were presence/absence of auditing, presence/absence of legal recourse, and the interaction of the two. Although the model was significant ($p<0.002$), the inter-

---

[22] The distributional assumptions made by an ANOVA cannot be met with the data (e.g., effort is bounded at zero and one). However, the number of observations removes some concern about the propriety of this parametric test. Where possible, nonparametric tests, including approximate randomization (Edgington 1980;

Table 6—*Continued*

Cell D Results
Auditing Available/Legal Recourse Available

Market (Audit Choice)	*Frequency of Legal Action*			
	Early		*Late*	
D1 (No)	100.0%	(2 of 2)	100.0%	(2 of 2)
(Yes)	—	—	—	—
D2 (No)	100.0	(2 of 2)	100.0	(2 of 2)
(Yes)	—	—	0.0	(0 of 1)
D3 (No)	100.0	(1 of 1)	—	—
(Yes)	—	—	—	—
D4 (No)	100.0	(1 of 1)	100.0	(2 of 2)
(Yes)	100.0	(1 of 1)	100.0	(1 of 1)
D5 (No)	100.0	(1 of 1)	—	—
(Yes)	—	—	—	—
D6 (No)	100.0	(2 of 2)	—	—
(Yes)	—	—	—	—
D7 (No)	—	—	100.0	(1 of 1)
(Yes)	—	—	—	—
D8 (No)	—	—	100.0	(1 of 1)
(Yes)	—	—	—	—
Average (No)	100.0%		100.0%	
Average (Yes)	100.0%		50.0%	
Average (All)	100.0%		92.9%	

Noreen 1989), were also performed. These tests produce the same conclusions as the parametric tests. Further, all ANOVAs used generalized least squares to control for unequal cell sizes for some tests. Logit, probit, and arc sine transformations were calculated, none of which changed the conclusions of any of the tests. In conformance with Stone and Rasp (1991), the sample size suggests least squares may be more appropriate and will be the method presented. Also, an investigation was made of the effects of removing an apparent outlier, market C6, the market with the lowest effort level. Market C6 had effort 3.1 standard deviations less than the mean of the other C markets (i.e., significantly lower, $p=0.0002$, two-tailed). Because statistically significant results become no less significant after dropping of C6, no change in the conclusions would be made, except as noted later in the test of net economic efficiency.

Table 7
Summary of Results

Market	*Effort**	*Truth***	*Market*	*Effort*	*Truth*	*Market*	*Effort*	*Truth*	*Market*	*Effort*	*Truth*
A1	33.3	44.4	B1	50.0	40.0	C1	92.3	80.0	D1	100.0	66.7
A2	33.3	28.6	B2	46.7	14.3	C2	69.2	66.7	D2	69.2	80.0
A3	61.5	0.0	B3	100.0	60.0	C3	50.0	85.7	D3	100.0	100.0
A4	91.7	77.8	B4	85.7	100.0	C4	53.8	42.9	D4	66.7	83.3
A5	38.5	87.5	B5	100.0	100.0	C5	92.3	100.0	D5	93.3	100.0
A6	20.0	71.4	B6	61.5	20.0	C6	69.2	57.1	D6	80.0	100.0
A7	22.2	71.4	B7	50.0	50.0	C7	35.7	100.0	D7	88.9	100.0
A8	18.2	88.9	B8	53.8	0.0	C8	41.7	100.0	D8	92.9	100.0
All A	39.8	58.8	*All B*	67.8	49.4	*All C*	51.0	79.0	*All D*	85.9	92.1
No Auditing No Legal Recourse			*Auditing No Legal Recourse*			*No Auditing Legal Recourse*			*Auditing Legal Recourse*		

* Percentage of high effort in late market days.
** Percentage of truthful disclosures in late market days.

action factor was not ($p=0.68$). The reduced model considering only main effects was significant ($p=0.0005$). In both the full and reduced models, the factors for auditing (both $p=0.0002$) and legal recourse (both $p \leq 0.04$) were significant in the predicted direction.

The results of the previous two ANOVAs suggest the following. First, both hypotheses H1a and H1b received strong support. That is, the presence of either legal recourse or auditing significantly increased effort levels (32 percent and 15 percent increases, respectively). However, with significantly positive increases for both auditing and legal recourse, the interaction factor in the full model should be significantly negative for hypothesis H1c to hold. Yet, it was positive and not significant. This supports a rejection of hypothesis H1c.[23] The presence of either auditing or legal recourse increased effort; the presence of both led to even higher effort.

The second series of hypotheses related to the level of audit purchases. Although hypothesis H2a is consistent with the analytical development, certain difficulties present themselves when a statistical test is considered. Since no auditing can be purchased in cell A, a single audit purchase in cell B is necessarily statistically significant. Further, any lack of audit purchase in cell B would cause the rejection of hypothesis H2a. Despite these problems, it is important to consider the proportion of audit purchases when no legal recourse exists.

The overall 50 percent level of audit purchases in cell B suggests a belief by the managers that auditing has value. Because the auditing system has a cost, there should be no incentive to purchase it unless it provides value (particularly by late market days). For this reason, the observed level was compared to a zero level of purchases. Since zero (and 100 percent) lies 2.3 standard deviations ($p=0.025$, one-tailed) from the observed 50 percent rate, it is reasonable to conclude that the demand for auditing is neither zero nor absolute. Simply, there is a demand for auditing.

Consistent with the analytical development, hypothesis H2b predicts that legal recourse should reduce the proportion of audit purchases. To test this, an ANOVA was constructed with audit purchase proportion in each market as the dependent variable and the presence/absence of legal recourse as the factor (giving 16 observations from cells with auditing, B and D). There was less frequent auditing in cell D (32 percent) than in cell B (50 percent) with mild significance ($p=0.08$). This allows a conclusion that auditing is purchased when available and that legal recourse drives out some, but not all, demand for auditing.

To test hypothesis H3a, a 1×3 ANOVA was constructed. The dependent variable was the proportion of high effort and the factor created three groups (containing eight observations each): cell A, cell B without audit purchase, and cell B with audit purchase. A Tukey comparison of group means showed that (1) effort was significantly higher with audits in cell B than in either cell A ($p<0.01$) or market days of cell B without audits ($p<0.05$), and (2) the effort in cell A and market days without audits in cell B did not differ significantly ($p>0.6$, two-tailed). Thus, there is significant evidence of increased effort with the purchase of auditing; hypothesis H3a is supported.

To test hypothesis H3b, a similar 1×3 ANOVA was constructed with effort as the dependent variable. The factor created three groups: cell C (eight observations), cell D without audit purchase (eight observations), and cell D with audit purchase (six observations). A Tukey comparison of group means showed that (1) effort was significantly higher for market days with audits in cell D than in cell C ($p<0.006$, two-tailed), (2) effort was not significantly different for market days with and without audits in cell D ($p>0.4$, two-tailed), and (3) the effort without audit in cell D was higher than in cell C, with mild significance ($p=0.06$, two-tailed). In sum, the results run contrary to hypothesis H3b. However, it appears that the availability of auditing rather than its actual purchase drives effort differentials when legal recourse exists. This confirms the results from the first series of hypotheses and will be discussed further shortly.

---

[23] This is a failure to reject the null hypothesis of no difference. An investigation of the power of this test suggests that eight markets per cell (with the observed variance) provide reasonable power for this test ($\beta \approx 0.05$).

The fourth hypothesis concerns the level of truthfulness. To test this, only market days with a low-quality asset were considered. A truthful disclosure was considered as a disclosure of "low" or "either." The proportion of truthfulness within each market was calculated and used as the dependent variable. The factors were presence/absence of auditing, presence/absence of legal recourse, and the interaction of the two. As predicted, the main and interactive effects involving auditing were not significant in the full and reduced (main-effects-only) models (both $p>0.25$). The reduced model with legal recourse as the sole factor showed that significantly more truthfulness occurred when legal recourse was present ($p=0.0006$). The addition of legal recourse increased truthfulness from 53 to 86 percent.[24]

The final series of hypotheses centers on the relative level of economic efficiency. To derive this, three terms were defined—maximum benefit (MB), gross benefit (GB), and net benefit (NB):

$$
\begin{aligned}
MB &= \Sigma e_1 \times HRV - C(e_1), \\
GB &= \Sigma e_i \times WRV - C(e_i), \text{ and} \\
NB &= \Sigma e_i \times WRV - C(e_i) - C(\nu_i) - C(l_i).
\end{aligned}
$$

The term e_i referred to the expected benefit from the selected effort level, where e_0 was 0.26 and e_1 was 0.74. HRV was the highest redemption value of the three buyers on a given day; WRV was the redemption value of the winning bidder. The cost operator, $C(\bullet)$, gives the cost of either selected effort, $C(e_i)$, auditing, $C(\nu_i)$, or legal costs, $C(l_i)$. The values of $C(e_1)$, $C(\nu_1)$, and $C(l_1)$ were 40, 30, and 50 cents, respectively, while $C(e_0)$, $C(\nu_0)$, and $C(l_0)$ were zero.

MB defines the maximum benefit of high effort to the economy, 74 percent of the highest redemption value net of the cost of high effort. GB defines the gross benefits actually received, net of actual effort cost. Even if the seller selected high effort every day, GB might be less than MB. This occurs when the winning bidder does not have the highest redemption value. Thus, the difference between GB and MB captures loss of economic efficiency through low effort or less-than-optimal allocation of resources (i.e., the "wrong" buyer purchases). NB was the GB value reduced by the auditing or legal costs actually incurred.

Gross economic efficiency was GB divided by MB. Net economic efficiency was NB divided by MB. If cooperation were possible without auditing or lawsuits, gross and net economic efficiency of 100 percent would be possible. If auditing were necessary for the expenditure of high effort, the maximum net economic efficiency with the selected

---

[24] Not surprisingly, the former is not different from a randomly (50 percent) truthful strategy ($p>0.9$) and the latter does not differ from the predicted 14 percent ($p=0.9$).

Table 8
Economic Efficiencies

Market	*Gross Economic Efficiency*	*Net Economic Efficiency*	*Market*	*Gross Economic Efficiency*	*Net Economic Efficiency*
A1	55.9	55.9	C1	93.9	91.2
A2	58.7	58.7	C2	76.7	71.9
A3	79.6	79.6	C3	60.9	60.9
A4	86.0	86.0	C4	79.2	74.1
A5	50.4	50.4	C5	72.4	72.4
A6	46.0	46.0	C6	37.5	27.8
A7	56.4	56.4	C7	66.9	66.9
A8	46.8	46.8	C8	60.1	60.1
A Average No Auditing/ No Legal Recourse	60.0	60.0	*C Average No Auditing/ Legal Recourse*	68.4	65.7
B1	74.1	67.5	D1	96.4	90.4
B2	68.3	55.5	D2	80.8	67.5
B3	97.1	95.6	D3	95.5	81.8
B4	92.4	81.8	D4	82.3	65.4
B5	100.0	85.2	D5	81.0	81.0
B6	71.3	61.1	D6	88.3	72.5
B7	72.1	62.4	D7	94.5	89.3
B8	78.0	63.6	D8	89.6	84.2
B Average Auditing/ No Legal Recourse	81.7	71.6	*D Average Auditing/ Legal Recourse*	88.5	79.0

parameters would be approximately 80 percent. An equilibrium of high effort in the environment with legal recourse would have net economic efficiency of approximately 95 percent.[25] Table 8 shows the level of gross and net economic efficiency for each market.

To test whether the level of economic efficiency differed among cells, two ANOVAs were constructed with either gross or net economic efficiency as the dependent variable. The factors were the presence/absence of auditing, the presence/absence of legal recourse, and the interaction of the two. Neither model showed significant interaction ($p > 0.5$ for both) and only main effects were considered. Like the ANOVA result for effort, this result causes a rejection of hypothesis H5c for gross efficiency. Although hypothesis H5c would be appropriate for gross economic efficiency, net economic efficiency should be higher when both auditing and legal recourse are available, compared with auditing only. However, a revised version of hypothesis H5c also would be rejected for net economic efficiency ($p > 0.5$).

---

[25] Note that the worst case has economic efficiency above zero. This measure was designed to be consistent with that in Dopuch et al. (1989).

The reduced model using gross efficiency was significant ($p=0.0002$). The presence of auditing ($p=0.0001$) and legal recourse ($p=0.05$) lead to higher gross economic efficiency. The reduced model using net economic efficiency was also significant ($p=0.03$) and showed that both the presence of auditing ($p=0.009$) and legal recourse ($p=0.10$) provided higher net economic efficiency. If experiment C6 (the apparent outlier noted previously) were removed, auditing would be significant at $p<0.0001$ and $p=0.017$ and legal recourse would be significant at $p=0.014$ and $p=0.024$, for gross and net efficiencies, respectively. This provides strong support for hypothesis H5a for both gross and net efficiency and mild (for net) to strong (for gross and for net with C6 removed) support for hypothesis H5b. Hypothesis H5c was not supported using either gross or net efficiency; auditing and legal recourse were complements, not substitutes, here.

The economic efficiency tests provide an interesting contrast to the similar ANOVA constructed with effort level as the dependent variable. The weakening of the legal recourse effect in the economic efficiency tests comes from a more frequent less-than-optimal allocation of the asset.[26] As with effort, the presence of either auditing or legal recourse increases economic efficiency, with the highest efficiency when both are present. It is interesting to note that the cell with the highest gross and net efficiency most closely matches naturally occurring markets.

Additional Observations

A further analysis was made of the buyers' bidding practices. Over- and underbidding can have important implications beyond the transfer of wealth between buyers and sellers. In general, it can lead to suboptimal resource allocations within an economy. Overall, the bid ratio should equal approximately the proportion that the lottery won. Comparisons were made for each market separately. Also, separate consideration was made of market days with and without audits in cells B and D. In all cells but one (including the split B and D cells), there was no significant over- or underbidding (all $p>0.17$). In cell A, the buyers significantly ($p=0.008$) overbid by approximately 33 percent.

An additional consideration is the integration of these results to those found in Dopuch et al. (1989). Cells A and B differed from two in Dopuch et al. (1989) in the time horizon assumption. Dopuch et al. used a known endpoint of 40 periods. Although the addition of an unknown horizon seems to make cooperation no less likely, effort levels in cell A were less than those in the corresponding cell in Dopuch et al., although the difference was not statistically significant ($p=0.2$, two-tailed).[27] When auditing was present, the markets reported here had essentially the same level of high effort ($p>0.9$, two-tailed). However, the 50 percent rate of audit purchase in cell B was less than the corresponding 77 percent rate in Dopuch et al., with mild significance ($p=0.07$, two-tailed). This provides support for the contention that, when the horizon is not known, the choice not to purchase auditing may signal a desire to reach the cooperative solution without incurring audit costs. No such signal was available in cell A.

---

[26] This is quite intuitive. On days when auditing is purchased, the buyers will generally know (or be able to deduce) actual quality. The uncertainty about actual quality when auditing is not purchased leads to more frequent occurrences of the highest bid by "wrong" buyers.

[27] The comparison is between a group of eight per-cell means from these markets and three per-cell means (from the late market days) from Dopuch et al. The small number of matching Dopuch et al. markets makes it difficult to find significant results.

Discussion

A final consideration relates to the considerably weaker effects of legal recourse on effort levels relative to auditing. While the analytic models indicate that either should work effectively but that legal recourse will drive out the demand for auditing, this was not observed. A possible explanation of the relative weakness of legal recourse here may come from ambiguity in the market. With an "agree" audit opinion, the investors are certain of asset quality; they are certain what probability distribution is used to determine the outcome. When legal recourse is present and auditing is not, the equilibrium prediction is that, when high quality is reported, it is uncertain which distribution will be used. The asset will be high (low) 96.2 (3.8) percent of the time and the asset will win with an 86.9 percent frequency. A growing body of literature (e.g., see Einhorn and Hogarth 1986, Hazen and Lee 1991) considers whether individuals avoid or seek ambiguous situations. Ambiguity-averse or -seeking individuals would value this asset differently from an asset that has a 100 percent chance of winning with an 86.9 percent frequency. There is also ambiguity on which value of λ will be used. This all works to add potential noise to the legal recourse environment that is not present in the auditing environment.

Another explanation may come from the complexity of the legal recourse equilibrium. Managers must follow a mixed strategy on lying (consistent with the data), investors must follow a mixed strategy on bringing suit (there were too few opportunities to determine if this occurred), and, from this, the manager can reap the benefits from high effort because all parties know this is in the manager's best interest. The "knife-edge" nature of this equilibrium might have made it difficult to achieve. The higher share price comes from an investor's belief in managerial action. This is not so with auditing.

The manager effectively creates an antifraud rule with the purchase of auditing. Therefore, investors will know asset quality with certainty. The investors need not believe the manager exerted high effort, they need only value the asset at its underlying value. The manager can get high values more often with high effort. Further, the manager can signal an intent to reach a first-best solution by not purchasing auditing (the second-best solution). However, this was not as effective in cell B, as shown by the lower effort when auditing was not purchased. This provides an advantage to the presence of legal recourse in cell D for managers seeking to reach a high-effort equilibrium without the purchase of auditing. Half the sellers in cell D had reasonable success with this strategy. Thus, results here suggest that legal recourse can help support somewhat higher levels of effort when the non-audit path is followed. This may be an important feature when auditing is not perfect. The difference is that auditing is primarily a forward-induction equilibrium, while legal recourse is a backward-induction equilibrium.

Finally, the result may come from a manager who seeks to act opportunistically. It is difficult to cheat on the auditing system, since the investors should be able to deduce actual quality when auditing has been purchased. However, when legal recourse is available, the manager should not be able to get away with opportunistic behavior in the long run, although it may be thought possible in the short run.

V. Summary and Conclusions

The results suggest that only a low level of cooperative behavior was possible when neither auditing nor legal recourse was available. This environment had the lowest level of seller effort and the only significant amount of overbidding by the buyers. The addition of either auditing or legal recourse led to higher levels of seller effort, the highest level coming when both options were available. The truthfulness of disclosures was higher when legal recourse was available. The tendency of buyers to overbid was reduced by both auditing and legal recourse. This suggests value to both the availability of auditing and legal recourse.

Overall, the results of this study suggest that the addition of auditing, legal recourse, or both leads to increased economic efficiency. This comes from the increased level of seller effort and the more efficient allocation of resources. However, the effect of legal recourse is somewhat small. Interestingly, the environment with the highest level of economic efficiency, where both auditing and legal recourse are available, is the one most nearly matching that currently found in securities markets.

There are several additional avenues for future research. One extension is the addition of mandated auditing. Under current legislation and SEC regulation, the presentation of audited financial statements is required in many instances. This may have a strong influence on the effect of (or even need for) antifraud rules on disclosures. Additional extensions would include modifications to the environment to allow a less-than-perfect legal system or a less-than-perfect auditing mechanism. The modification of the auditing mechanism may come about through random error or moral hazard. Since there is often a time lag between the investors' loss and recovery, the discounting of damages is another possible extension. The legal system could then be extended to include suits against the auditor. Also, the effect of the form of the legal system (e.g., American, British, treble damages, etc.) could be investigated.

This research may provide some early feedback for the development of theories on the demand for auditing. The results presented here suggest that the legal system should be considered in the modeling of audit demand. This line of research also has implications for the modeling of the regulation of securities markets and disclosures. In the longer term, this line of research could possess important public policy implications.

References

Axelrod, R. 1984. *The Evolution of Cooperation.* New York: Basic Books.

Baiman, S. 1979. Discussion of "Auditing: incentives and truthful reporting." *Journal of Accounting Research* 17 (Supplement): 25–9.

———, J. Evans, and J. Noel. 1987. Optimal contracts with a utility-maximizing auditor. *Journal of Accounting Research* 25 (Autumn): 217–44.

Bendor, J. 1986. Tit-for-tat in the noisy prisoner's dilemma. Working paper, Graduate School of Business, Stanford University, Stanford, CA.

———. 1987. In good times and bad: reciprocity in an uncertain world. *American Journal of Political Science* 89 (August): 531–58.

Blazenko, G., and W. Scott. 1986. A model of standard setting in auditing. *Contemporary Accounting Research* 3 (Fall): 68–92.

Cox, J., B. Roberson, and V. Smith. 1982. Theory and behavior of single object auctions. In *Research in Experimental Economics,* vol. 2, edited by V. Smith. Greenwich, CT: JAI Press.

DeJong, D., R. Forsythe, R. Lundholm, and W. Uecker. 1985a. A laboratory investigation of the moral hazard problem in an agency relationship. *Journal of Accounting Research* 23 (Supplement): 81–120.

———, ———, and W. Uecker. 1985b. The methodology of laboratory markets and its implications for agency research in accounting and auditing. *Journal of Accounting Research* 23 (Autumn): 753–93.

Dopuch, N., and R. King. 1991. *Experimental Tests of Auditing as a Credibility Generating Mechanism.* Vancouver, British Columbia, Canada: Canadian Certified General Accountants Research Foundation.

———, and ———. 1992. Negligence versus strict liability regimes in auditing: An experimental investigation. *The Accounting Review* 67 (January): 97–120.

———, ———, and D. Wallin. 1989. The use of experimental markets in auditing research: Some initial findings. *Auditing: A Journal of Practice & Theory* 8 (Supplement): 98–127.

Dye, R. 1985. Disclosure of nonproprietary information. *Journal of Accounting Research* 23 (Spring): 123–45.

———. 1990. Mandatory versus voluntary disclosures: The cases of financial and real externalities. *The Accounting Review* 65 (January): 1–24.

Edgington, E. 1980. *Randomization Tests.* New York: Marcell Dekker.

Einhorn, H., and R. Hogarth. 1986. Decision making under ambiguity. *Journal of Business* 59 (Supplement): S225–50.

Evans, J. 1980. Optimal contracts with costly conditional auditing. *Journal of Accounting Research* 18 (Supplement): 108–28.

Forsythe, R., R. Isaac, and T. Palfrey. 1989. Theories and tests of "blind bidding" in sealed-bid auctions. *Rand Journal of Economics* 20 (Summer): 214–38.

Grossman, S. 1981. The informational role of warranties and private disclosure about product quality. *Journal of Law and Economics* 24 (December): 461–83.

———, and O. Hart. 1980. Disclosure laws and takeover bids. *Journal of Finance* 35 (May): 323–34.

Hazen, G., and J. Lee. 1991. Ambiguity aversion in the small and in the large for weighted linear utility. *Journal of Risk and Uncertainty* 4 (April): 177–212.

Isaac, R., and J. Walker. 1986. Communication and free riding behavior: The voluntary contributions mechanism. Working paper, College of Business and Public Administration, University of Arizona, Tucson.

Jensen, M., and W. Meckling. 1976. Theory of the firm: Managerial behavior, agency costs and ownership structure. *Journal of Financial Economics* 3 (October): 305–60.

Kachelmeier, S. 1991. A laboratory market investigation of the demand for strategic auditing. *Auditing: A Journal of Practice & Theory* 10 (forthcoming).

King, R., and D. Wallin. 1990a. The effects of antifraud rules and *ex post* verifiability on managerial disclosures. *Contemporary Accounting Research* 6 (Spring): 859–92.

———, and ———. 1990b. Optimal level of fraudulent disclosure when litigation is costly. Working paper, College of Business, The Ohio State University, Columbus.

———, and ———. 1991. Voluntary disclosures when seller's level of information is unknown. *Journal of Accounting Research* 29 (Spring): 96–108.

Klein, B., and K. Leffler. 1981. The role of market forces in assuring contractual performances. *Journal of Political Economy* 89 (August): 615–41.

Milgrom, P. 1981. Good news and bad news: Representation theorems and applications. *Bell Journal of Economics* 12 (Autumn): 380–91.

———, and J. Roberts. 1986. Relying on the information of interested parties. *Rand Journal of Economics* 17 (Spring): 18–32.

Murnighan, J., and A. Roth. 1983. Expected continued play in prisoner's dilemma games. *Journal of Conflict Resolution* 27 (June): 279–300.

Noreen, E. 1989. *Computer-Intensive Methods for Testing Hypothesis: An Introduction.* New York: John Wiley & Sons.

Ordeshook, P. 1986. *Game Theory and Political Theory.* Cambridge, United Kingdom: Cambridge University Press.

Roth, A., and J. Murnighan. 1978. Equilibrium behavior and repeated play of the prisoner's dilemma. *Journal of Mathematical Psychology* 17 (April): 189–98.

Scott, W. 1984. The state of the art of academic research in auditing. *Journal of Accounting Literature* 3: 153–200.

Stone, M., and J. Rasp. 1991. Tradeoffs in the choice between logit and OLS for accounting choice studies. *The Accounting Review* 66 (January): 170–87.

Townsend, R. 1982. Optimal multiperiod contracts and the gain from enduring relationships under private information. *Journal of Political Economy* 90 (December): 1166–86.

Wallin, D. 1990. Alternative economic institutions to motivate managerial disclosure of private information: An experimental markets examination. Ph.D. dissertation, University of Arizona, Tucson.

THE ACCOUNTING REVIEW
Vol. 67, No. 1
January 1992
pp. 148–156

A Perspective on the Use of Laboratory Market Experimentation in Auditing Research

James R. Boatsman
Lawrence P. Grasso
*Michael B. Ormiston**
J. Hal Reneau
and The Arizona State University Research Workshop

VERIFICATION of accounting information provided by managers of firms is essential to the efficient operation of a market economy. It is therefore important to understand the interaction of buyers and sellers in the market for auditing services, and how the operation of this market affects the economy as a whole.

The dominant theoretical model used to analyze auditing as a verification mechanism is based on the methodology developed in the principal-agent and game-theory literatures. In the basic model, there exists a single firm operated by a risk-neutral, effort-averse manager and owned by many risk-neutral investors. Owners and managers are assumed to be expected-utility maximizers. At the beginning of each period, the manager decides whether to purchase auditing and selects a level of effort. Effort, in turn, affects the firm's expected output for the period. The manager then receives private information concerning the firm's expected output, decides whether to disclose the information, and decides whether to release any audit report concerning expected output. Next, owners bid competitively for shares of the firm, thus establishing the market value of the firm. Finally, the firm's actual output is revealed to all market participants. At this point, owners may seek legal recourse from the manager if the disclosure is believed to have been fraudulent, or from the auditor if the audit report is believed to have been fraudulent. This sequence of events is usually modeled as a principal-agent problem, a repeated-play game with a known finite horizon, or a repeated-play game with an unknown or infinite horizon.

Various forms of this basic model have been used to address a variety of interesting issues concerning the market for auditing services. Most of the recent theoretical and empirical research, however, has been concerned with one of two issues: either the

---

*Michael B. Ormiston is in the Economics Department at ASU.

Commissioned by the Editor.
Received September 1991.

determinants of the demand for auditing services or the effects of alternative liability regimes. For example, Evans (1980), Antle (1982), and Baiman et al. (1987) investigate the nature and feasibility of contracts among owners, managers, and auditors to improve economic efficiency (i.e., contracts that internalize moral-hazard problems). The primary implication of this line of research is that a demand for auditing arises as part of the optimal contracting scheme. Dye (1985, 1990), King and Wallin (1990, 1991), Kachelmeier (1991), Dopuch et al. (1989), and the studies in this forum investigate the effects of antifraud rules, ex post verifiability, and legal liability rules on a manager's disclosure of private information.

Both of the forum studies develop and test a model in which auditing is available as a means of verifying a manager's disclosure of private information. Both incorporate legal recourse in some fashion. There are, however, some important differences. The primary goal of the Wallin study is to explain effects of competing forms of verification (auditing and legal recourse against *managers*). The primary goal of the Dopuch and King study is to explain effects of competing forms of legal recourse against *auditors* (negligence and strict liability). Dopuch and King incorporate competition among auditors, such that the price of auditing is endogenous. If hired, their auditors decide whether to undertake a costly test of the manager's disclosure. Deciding to test is necessary to preserve the option of disagreeing with the manager's disclosure. Wallin excludes competition among auditors; accordingly, the price of auditing is exogenous. Moreover, there are no counterparts of Dopuch and King's auditor testing and reporting decisions. Both articles develop a number of hypotheses, and both test the hypotheses with data obtained from experimental markets.

The remainder of this perspective is organized as follows. In section I we discuss some concerns regarding the experiments. Section II contains additional concerns regarding generality. A final section discusses the relative merits of game theory and more conventional economic analysis as means of characterizing equilibrium in the market for auditing services. In addition, we raise the question of whether *any* equilibrium characterization offers promise of dealing with many of the more important questions in auditing.

I. Concerns Regarding the Experiments

Although it is unreasonable to expect empirical results to conform perfectly to theoretical predictions, many of the experimental results discussed in the forum articles deviate substantially from the maintained hypotheses. In Wallin's study, for example, high effort should not be observed in cell A (no auditing or legal recourse available), auditing should not be purchased in cell D (both auditing and legal recourse available), and auditing should always be purchased in cell B (only auditing available). Nonetheless, high effort was observed in about 40 percent of the cell A trials, auditing was purchased in about a third of the cell D trials, and auditing was not purchased in about 50 percent of the cell B trials.

Deviant results can also be seen in the Dopuch and King data. For example, zero auditor hires should be observed in the no-liability markets. Yet auditors were hired in 61 percent of the trials, a number not reliably different from the 78 percent auditor hire rate observed in the negligence markets, where auditors should always be hired. Similarly, accepted audit offers averaged $0.28 in the no-liability markets although they should have been zero. The costly investment should never (always) be selected in the no-liability (strict liability) regimes, but was selected 45 (68) percent of the time. Qualitatively, we believe the empirical results deviate from the theoretical predictions enough to pose the question, "What can one infer when the prediction(s) tested in a laboratory market do not attain?"

Laboratory markets can be useful for testing a theory's predictions in a stylized setting where certain precepts are met. These precepts include nonsatiation, saliency, dominance, and privacy, and are sufficient conditions for a valid and controlled laboratory market (Smith 1982). When the observed experimental data are inconsistent with the predictions, the underlying reason could be that the theory is incorrect, that the precepts were not met, or both. Unfortunately, it is difficult to disentangle these explanations, given negative results.

Failure to satisfy the precepts could be manifest either as a failure to converge over time to an equilibrium (if a unique equilibrium exists) or as a failure of the market participants to implement some consistent game strategy. Both reasons could be expected to produce data with high variance and data that do not converge over time. The results of both studies show evidence of a lack of convergence and high variance.[1]

Dopuch and King acknowledge several reasons why their results might differ from the maintained hypotheses. These included (1) questions about the risk neutrality of the market participants, (2) the nature of the incentive provided in the second-price, sealed-bid institution, (3) the possibility of participants' attending to the multiple periods of the game, and (4) the possible influence of the insurance function in the strict liability market. They conclude that the participants appear to have adopted complex strategies and that the verifiers had difficulty pricing their service. In a similar vein, Wallin acknowledges the potential influence of complexity of the equilibrium under legal recourse, as well as the potential effect of ambiguity (which was not considered in the experiment).[2]

We agree with the authors' plausible reasons for why the data may differ from the predictions. However, we believe there are at least two additional reasons that are equally plausible and more problematic: the existence of multiple equlibria and the complexity of the game.

To his credit, Wallin incorporated into his experiments an aspect of uncertainty that exists in the real-world market and that had not been incorporated previously. He postulated markets with uncertain end points, yet his maintained hypotheses are based on the assumption that the single-period game predictions will hold even with an uncertain horizon. This seems contradictory at best, and is not generally true in repeated games with unknown horizons. Even in the simplest repeated games with unknown horizons, the set of feasible equilibria is quite large. Of course, the single-period equilibria are contained in this set, but there is no theoretical reason to suppose one of them obtains as the outcome of the game.[3] As a result, no definite theoretical predictions can be made (or empirical test devised) concerning behavior without making some *ad hoc* assumption about the heuristic used by players of the game. This is, of course, tantamount to assuming which equilibrium obtains at the outset. Unless some reason is given for the *ad hoc* assumption, nothing is gained in terms of explaining the behavior of buyers and sellers of audit services. Dopuch and King's predictions do not suffer from this problem since a certain horizon market was used. However, they speculate that their subjects may have used multiperiod strategies, such as those that would be appropriate in an uncertain horizon game. If their subjects did view the game as if the horizon were uncertain, their predictions are questionable.

---

[1] Refer to figure 2 and tables 3–6 in Wallin and tables 3 and 4 in Dopuch and King. Further, both studies provide evidence that the conclusions reached from analysis of only the late market periods do not substantively differ from the conclusions reached when analyzing all periods. This is consistent with high-variance data that do not converge.

[2] For a discussion of the effects of ambiguity in insurance decisions, see Hogarth and Kunreuther (1989).

[3] Wallin provides an extensive discussion of multiple equilibria in a multiperiod game with an uncertain end period. He also cites some evidence supporting his conjecture that the noncooperative behavior predicted

To illustrate that the games may simply be too complex for subjects to understand, consider the buyers and verifiers in Dopuch and King's negligence regime. Rational buyers and verifiers would need to solve two sets of three simultaneous equations, one set for each possible investment type. Next, they must determine the seller's strategy with the highest expected profit to the seller. The solution in this case results in mixed strategies for both buyers and verifiers: verifiers shirk in about 9 percent of the trials and buyers sue in about half.

It is not necessary for the market participants to actually be the rational players described above. The predicted equilibrium can be reached because the market participants learn from their experiences over the game's trials. Learning, however, becomes more difficult when uncertainty increases (i.e., when one faces a compound probability assessment for outcomes and players follow complex strategies).[4]

The game structures in both forum studies are very similar. We characterize the games as complex and as having a high degree of uncertainty. While we commend the authors' substantial efforts directed at enhancing subjects' learning, we fear these efforts may have been insufficient.

II. Concerns Regarding Generality

Modeling the market for auditing services as a repeated-play game presents a dilemma. On one hand, it allows one to show that a demand for auditing exists because verification of disclosures is a means of mitigating the inefficiencies caused by information asymmetries. But on the other hand, it forces one to sacrifice many potentially important aspects of the real-world market for auditing services in order to maintain tractability. This is not to say that a theoretical model should be judged on the basis of whether it resembles the appallingly complex real world. Rather, its usefulness depends on whether it succeeds in explaining and predicting the phenomena of interest. By its very nature, a theoretical model should simplify by abstracting the most important factors of the problem at hand. Within the confines of a game theoretic formulation of the market for auditing services, however, simplifying for the purpose of maintaining tractability may exorcise the heart of the phenomena of interest. Both forum studies have simplified or ignored many important aspects of the market for au-

---

for the single-period, no auditing/no legal recourse scenario would persist in a multiperiod setting. However, he provides no discussion of why the single-period equilibria predicted for the other three scenarios (auditing available without legal recourse, legal recourse available without auditing, and both auditing and legal recourse available) would be expected in a multiperiod setting.

[4] Consider the difficulty in learning from observed outcomes over 20 periods when the outcome is a joint function of your opponent's choice and uncertainty in the payoff distribution; further, your opponent plays a mixed decision strategy with choices that are approximately equiprobable.

diting services to maintain tractability, and this limits the robustness of their theoretical predictions and empirical tests.

To illustrate this point, consider first the matter of reputation. Few would question the importance of reputation in the market for auditing services. But as both Wallin and Dopuch and King discuss, there is no theoretical reason for reputation-building strategies when all players know with certainty when the last play of the game will occur. In an attempt to capture this important element of the real world in his experiment, Wallin's subjects were precluded from knowing which market period would be the last. As noted above, incorporating this feature introduces a multitude of equilibrium strategies and confounds theoretical predictions as well as empirical tests.

Dopuch and King avoid the problem of multiple equilibria by confining the model to a finite number of market periods (20). Although this enables them to develop unique testable hypotheses, it strips from their model the theoretical reasons for reputation building. This is particularly problematic in their strict liability scenario, where verifiers provide an insurance service in addition to an attestation service. The only way verifiers can limit their exposure as insurers is by testing. Testing preserves the option to disagree with false seller disclosures, which dissuades purchasers from paying too much for a low-quality asset. But in the real world, insurers employ other means of limiting their exposure. For example, writers of auto insurance contracts routinely investigate driving records prior to providing (or continuing to provide) coverage. A single price of auto insurance doesn't exist. Poor drivers pay more than their more careful counterparts. Knowing this, insured drivers sometimes avoid processing valid claims in order to preserve a reputation as a good driver. One can certainly envision how verifiers might limit their liability exposure by investigating a seller's investment history instead of by extensive current testing. One real-world counterpart of this sort of trade-off is the assessment of inherent risk in choosing a level of substantive testing. Similarly, sellers of equity claims have an incentive to employ verifiers with a reputation for quality verification.

Continuing with this theme, we note that both the Wallin and Dopuch and King market settings involve complete settling up at the end of each market period. In each subsequent period, all players continue their roles. In the real world, however, owner/managers become employees in periods after the one in which they sell the initial equity claim to their firm. At this point, holders of equity claims can write incentive contracts to induce risk-averse managers to exert high levels of effort. It is easy to envision how incentive contracts might displace the demand for verification that exists in the Wallin and Dopuch and King market settings. This is not to say, however, that a demand for verification might cease to exist. In the real world, there is a secondary market for equity claims. Accordingly, current owners of high-quality assets have an incentive to submit assertions about historical payoffs to verification. A market for debt introduces a demand for verification as well.

Another important real-world factor absent from game theoretic formulations in general, and the forum studies in particular, is the effect of market structure on behavior. For example, in Dopuch and King's markets the cost of verification is $0.50 regardless of whether the verifier is an incumbent. This is quite reasonable since the two sellers are engaged in identical activities. In the real world, however, the counterparts of Dopuch and King's sellers are not engaged in identical activities. For this reason, initial audits are more costly than subsequent audits. DeAngelo (1981a) has shown that this fact, coupled with competition among auditors and costs of switching auditors, leads to low balling on initial engagements and an ability to earn economic rents on repeat engagements. The series of future rents is sacrificed upon losing a client. If auditors had only a single client, the prospect of such a sacrifice would create an incentive to acquiesce to client misrepresentation. But auditors have many clients, and acquiescing to misrepresentation by any one client carries the prospect of sacrificing the economic rents that would otherwise be earned on many other clients. Accordingly, DeAngelo (1981b) argues that auditors with many clients have strong incentives to perform diligently. In short, the structure of the real-world market for auditing services implies potentially much harsher penalties for deviant behavior than is characterized in either of Dopuch and King's liability scenarios.

In Wallin's markets, the harshness of penalties is crucial to his single-period prediction that legal recourse drives out the demand for auditing when both are available. In fact, this prediction is misstated. An auditing activity occurs in both the auditing available/no legal recourse and the no auditing/legal recourse available scenarios. In the former, managers can contract for an audit prior to choosing to exert effort. In the latter, purchasers can contract for an audit upon observing an unfavorable outcome. Although Wallin doesn't speak of a purchaser's decision to sue as contracting for an audit, this is what occurs. Recall that what he terms auditing and legal recourse are perfect substitutes in the sense that both determine asset quality with certainty. Thus, a more precise statement of his prediction is that managers will relegate the audit decision to purchasers when it is possible to do so. But this is dependent on the particular numerical values chosen. Wallin's litigation cost is not "much larger" than the cost of auditing. Accordingly, the manager who seeks to maximize expected value will avoid the certain cost of contracting for an audit, favoring instead the lower expected cost of legal recourse.

It is troublesome that Wallin's prediction depends on particular numerical values, especially when one can envision real-world scenarios in which the expected cost of litigation exceeds the certain cost incurred when a manager contracts for an audit. Consider, for example, superimposing punitive damages on Wallin's penalty structure. High punitive damages can make litigation cost "much larger" than the cost of auditing and reverse his prediction.

In the markets of both Dopuch and King and Wallin, the players are endowed with extensive information. Purchasers, for example, know that sellers can choose between a costly and a costless investment. They know the distributions of high- and low-quality assets associated with each choice, and the distributions of high and low payoffs associated with each asset type. In short, purchasers have the requisite information to determine a seller's strategy (at least in a single-period setting). Without assuming that purchasers had such information, it would be impossible to generate behavorial predictions. In the real world, however, purchasers are not so well informed. Indeed, one of the presumed reasons for separation of ownership and management is that managers enjoy some comparative advantage in identifying and evaluating investment opportunities.

Predictions of economic behavior almost always demand seemingly outrageous assumptions. Whether assumptions actually hold in the real world is normally of little concern. At best, economic theories are *as if* characterizations of behavior. If a theory predicts well, the customary interpretation is that market agents behave *as if* the assumptions hold. The problem with the two experiments at hand is that the assumptions regarding what each player knows are *allowed* to hold. Experimental outcomes that confirm predictions may arise (1) because the informational assumptions are allowed to hold or (2) because subjects behave as if they have the requisite information to implement complex strategies. The distinction is important. Generalizing to a world in which the assumptions don't hold is considerably more limited if experimental outcomes arise because informational assumptions are allowed to hold. Unfortunately, the experimenters can't tell the difference.

In Dopuch and King's markets, negligence is operationalized as a conscious choice to shirk. Liability is predicted to motivate verifiers to avoid shirking. Nonshirking, in turn, motivates sellers to select the costly investment. And the investment choice ultimately affects the level of wealth generated in the economy. In the real world, however, outright shirking is essentially impossible because of peer review and intrafirm control procedures. A variety of auditor actions are deemed negligent, but outright shirking, or even underinvesting in test procedures, are rarely among them. Penalties for negligence probably do affect auditor behavior. But it is far from clear whether the behavioral effect is one that influences a client's ranking of alternative investments. Indeed, the level of auditor testing plays no role in conventional capital budgeting models, but a theoretical linkage between auditor liability and investment policy is critical to Dopuch and King's predictions about wealth creation.

The above litany could continue *ad nauseam*. Neither audits nor legal proceedings are perfectly revealing. There is no real-world counterpart of a "disagree" auditor opinion accompanying a registration of securities in accordance with the Securities Act of 1933. Judicial remedies affect issuers and underwriters as well as auditors. Issuers cannot elect to suppress an audit opinion. The point is that both of the models at hand abstract from the real world in ways that can affect predictions. Thus, even if the experimental outcomes had confirmed all predictions, a large number of grounds upon which to question their generality would persist. The seriousness with which one questions is probably related to the extent to which one subscribes to the game theoretic paradigm. Skeptics are inclined never to be convinced that a meaningful level of realism has been achieved. True believers don't much care.

III. Alternative Approaches to Characterizing Equilibrium

The game theoretic formulation used to generate testable hypotheses provides interesting insights into two important aspects of the market for auditing services: (1) the importance of auditing as a means for internalizing the moral-hazard problem and (2) the importance of legal liability rules for determining the equilibrium price and quantity of auditing services. However, as noted above, game theoretic formulations suffer serious deficiencies that bring into doubt the robustness of theoretical and empirical results. Further, attempts to correct for these deficiencies will introduce sufficient complexities to render a game theoretic formulation intractable.

Fortunately, game theory is not the only type of equilibrium analysis used in economics to analyze decision making under uncertainty and associated market behavior. For example, the economics literature abounds with models of insurance markets where competitive pressures render strategic behavior relatively unimportant. In these models, buyers and sellers are assumed to be expected-utility maximizers and the market for insurance is assumed to be competitive. Determining equilibrium prices and quantities is a relatively straightforward task, as is analyzing the effects of alternative market structures, liability rules, and pricing schemes.

The market for auditing services could be modeled in a similar fashion. For example, one might model (from an investor's point of view) auditing as insurance against fraudulent disclosure. In such a model, each investor chooses the amount of auditing that maximizes expected utility associated with having a fixed proportion of wealth invested in a given firm. Expected utility depends on the level of auditing services purchased and the probability distribution characterizing the uncertain rate of return. The rate of return has two sources of uncertainty: uncertainty regarding the state of nature and uncertainty regarding fraudulent disclosure. A demand for auditing arises in such a model because auditing provides investors with valuable information concerning the quality of the firm. Moreover, a competitive market for auditing services obviates concern for strategic behavior.

The cost of using such a model to analyze the market for auditing services appears small in relation to the gain. One sacrifices the ability to explain precisely the nature and feasibility of contracts designed to internalize moral-hazard problems (although, in fact, one can still discuss such issues in a broader fashion). However, one gains the ability to address many of the issues raised in this perspective while maintaining tractability.

Note that proceeding in such a fashion is in no way inconsistent with the use of experimental markets to test predictions. But because strategic behavior is unimportant, it is likely that the market settings would be simplified. In contrast, the laboratory markets settings in the forum studies are complex by experimental economics standards despite the exclusion of potentially important real-world attributes. In settings where strategic behavior is important, a generalizable theory of auditing is unlikely to emerge until a vast number of experiments have been conducted.

Regardless of whether game theory or more conventional economic analysis is used, focus on the existence, uniqueness, stability, and efficiency of equilibrium within the context of a given set of institutions may not be useful for addressing some important auditing questions. Indeed, many of what we view as important questions in auditing are not obviously amenable to equilibrium analysis. One such question involves why auditing has evolved into its current institutional form. Why, for example, is auditing predominantly a mandated private sector (as opposed to a public sector) activity? Equilibrium analysis has not proved particularly useful in answering questions involving comparative institutional forms.

Another question of considerable interest involves scope of services. Are current levels of consulting activities high enough to degrade significantly the perceived credibility of the audit report? It is hard to imagine how an equilibrium analysis can determine whether current consulting levels are, in fact, high enough to have a significant deleterious effect. Yet another question is whether the demand for auditing is a demand for the detection of fraud, detection of corrupt business practices, or detection of disconformity with GAAP. Equilibrium analysis is unlikely to provide answers to questions regarding what sort of verification is demanded. Finally, a host of important questions about the effectiveness of audit technologies have little to do with equilibrium in the market for audit services.

The generally negative tone of this perspective should not be misinterpreted. Laboratory market studies have a clear potential for embellishing our understanding of auditing. Our hope is that we have at least stimulated readers to consider alternatives to the principal-agent and game-theoretic approaches for characterizing the market for audit services.

References

Antle, R. 1982. The auditor as an economic agent. *Journal of Accounting Research* 20 (Autumn): 503–27.

Baiman, S., J. Evans, and J. Noel. 1987. Optimal contracts with a utility-maximizing auditor. *Journal of Accounting Research* 25 (Autumn): 217–44.

DeAngelo, L. 1981a. Auditor independence, "low balling," and disclosure regulation. *Journal of Accounting & Economics* 3 (August): 113–27.

———. 1981b. Auditor size and audit quality. *Journal of Accounting & Economics* 3 (December): 183–99.

Dopuch, N., R. King, and D. Wallin. 1989. The use of experimental markets in auditing research: Some initial findings. *Auditing: A Journal of Practice & Theory* 8 (Supplement): 98–127.

Dye, R. 1985. Disclosure of nonproprietary information. *Journal of Accounting Research* 23 (Spring): 123–45.

———. 1990. Mandatory versus voluntary disclosures: The cases of financial and real externalities. *The Accounting Review* 65 (January): 1–24.

Evans, J. 1980. Optimal contracts with costly conditional auditing. *Journal of Accounting Research* 18 (Supplement): 108–28.

Hogarth, R., and H. Kunreuther. 1989. Risk, ambiguity, and insurance. *Journal of Risk and Uncertainty* 2 (April): 5–35.

Kachelmeier, S. 1991. A laboratory market investigation of the demand for strategic auditing. *Auditing: A Journal of Practice & Theory* 10 (Supplement): forthcoming.

King, R., and D. Wallin. 1990. The effects of antifraud rules and *ex post* verifiability on managerial disclosures. *Contemporary Accounting Research* 6 (Spring): 855–87.

———, and ———. 1991. Voluntary disclosure when seller's level of information is unknown. *Journal of Accounting Research* 29 (Spring): 96–108.

Smith, V. 1982. Microeconomic systems as an experimental science. *American Economic Review* 72 (December): 923–55.

THE ACCOUNTING REVIEW
Vol. 67, No. 1
January 1992
pp. 157–170

A Perspective on the Use of Laboratory Market Experimentation in Auditing Research

Douglas V. DeJong
Robert Forsythe
and The University of Iowa Accounting Workshop

EXPERIMENTAL methods have contributed to our understanding of economics over the last 30 years and have been used to address economic issues in auditing for more than five years. The methods of experimental economics offer several distinct advantages: they provide a high degree of control over an economic setting and the incentives of agents in these settings, and they allow the study of economic settings that do not exist or are not observable in the real world.

In this perspective, we discuss the use of these methods to address the role of experimental economics research in auditing and discuss the features of a good economic experiment. We take the view that the role of this research is to facilitate a dialog with theory about economic phenomena of interest in auditing. Economic theory generates testable hypotheses that drive an experimental design, and experimental data provide observations necessary to test theory. As a result of this dialog, theories are modified, refined, or discarded. A good experiment not only provides sufficient control to test a theory but also provides sharp insights about how the theory may be modified when the data contradict the theory's predictions. Any good experiment will almost always generate some contradictory evidence, and this should be useful in modifying the theory to explain the observed phenomena. Since our understanding of a theory is almost always incomplete, this kind of testing requires that the theory be tested as close to its domain as experimental control will allow.[1]

---

[1] It is often not possible to test a theory directly on its domain in the laboratory. For example, most economic theories require the assumption of common knowledge of agent rationality, and we currently know of no way to impose that assumption in a laboratory environment. (See Smith et al. [1988] for a discussion of what can happen if this assumption is violated.)

We would like to acknowledge the contribution of the workshop participants at the University of Iowa who did an outstanding job of analyzing the Forum papers. DeJong appreciates the support provided by the National Science Foundation and the McGladrey Institute of Accounting Research. Forsythe acknowledges the support provided by the National Science Foundation.

Invited by the Editor.
Received October 1991.

An alternative approach is to compare the performance of competing theories in some common setting (i.e., to test these theories off their domain). Assuming that the setting captures the essential details of some real-world setting, the researcher expects more than one theory should hold. This typically leads to a horse race between competing theories, which is often helpful in gaining some initial impressions of the theories. But, data on the relative (as opposed to absolute) performance of theories is less compelling than direct tests of each theory.

It is also argued that experiments can be useful in providing empirical regularities or "stylized facts" in settings where theory is silent due to intractability (see, e.g., Smith et al. 1987). Without a theory as an organizing principle, however, it is not clear what elements of the experiment must be controlled. For example, in the absence of theory, it would be difficult to know why the results of one experimental session differ from another when the levels of the "controlled" variables remained unchanged.

I. The Methodology of Experimental Economics

Smith (1976, 1982) and Plott (1982) were the first to detail the necessary steps to construct a microeconomic system in the laboratory, and others have provided additional information about the control and conduct of economic experiments (see, e.g., Berg et al. 1990; Cox and Isaac 1986; Smith et al. 1987). The principal point that we want to emphasize from this literature is that it is possible to create a real economic system with real economic agents in the laboratory. There is sufficient control over the environment (set of agents, set of commodities, resources, and agent preferences), institutions (property rights, rules of communication and exchange among agents), and agent behavior (choices made by agents from the set of allowable choices) such that economic theories can be tested.

In the laboratory, predictions can be tested in settings that come as close as possible to satisfying the assumptions of the theory that generated them. Although success at this level does not mean that the theory will be useful in naturally occurring settings, laboratory testing is an important and necessary step to determine whether a theory is valid on its own terms. If the theory is not valid on its own domain, it is unlikely to hold off its domain. Only after a theory has been fine-tuned is it interesting to use experimental methods to map out its boundary conditions by varying parameters and relaxing its assumptions.

Laboratory settings can be simple or complex. When theoretical predictions are rejected in a complicated setting, identifying the source of the failure can be difficult. In contrast, the source of a theory's failure in a simple setting can be more readily identified and the necessary theoretical modifications made. For this reason, we argue for simple settings in initial investigations. Examples that support this argument are presented in the next section, prior to our consideration of the Forum studies.

Testing Complex Settings

The experiment in DeJong et al. (1985) was designed to test the predictions of a single-period, competitive "lemons model" in the setting described below. The model assumes that the preferences of all agents, the sellers' costs, and the possible quality levels of services are commonly known by all market participants. In this model, a seller enters into a contract with a buyer to provide a service at a particular price. The quality of the service is not contractually guaranteed, nor could it be since, even after taking delivery, the buyer never directly observes the quality of the service supplied. The competitive "lemons" prediction is that, in equilibrium, all sellers will supply the lowest quality level and the market price will be equal to the cost of supplying that level of service. The model is silent about all other details of the trading institution.

The setting consisted of three sellers of a service and four buyers, who played a repeated game against one another. The service purchased from the sellers reduced the likelihood that buyers would face a potential loss where that likelihood depended upon the quality of service the seller supplied. In each period, the sellers posted individual prices (one for each buyer) and disclosed the quality of service they would supply. Each buyer then chose which seller, if any, to purchase from and, after the contract was made, sellers selected an actual level of service to supply, which did not have to be the same as the disclosed level. Further, a seller's cost increased with the quality supplied. After taking delivery, a buyer observed only whether a loss occurred but did not learn the actual level supplied. Profits were then calculated by each of the agents, and everyone proceeded to the next period, which was a strict replication of the previous period(s).

All agents knew they would keep the same roles (buyer or seller) throughout the experiment and no one knew in advance when the experiment would end. All agents knew the number of buyers and sellers in the market, the amount of the potential loss and the effect that the different quality levels had on this likelihood. Sellers knew only the cost structures for their own service levels, and buyers did not know the cost structures for any of the sellers. Agents had no direct knowledge of the preferences of other agents. Finally, agents knew only their individual sales, outcome histories, and profits. All of the above information was contained in instructions that were read aloud to all subjects at the beginning of the session. Subjects were all in the same room when the instructions were read; thus, this information should be common knowledge.

The data show that, although a large number of low quality goods were traded, the market failed to completely collapse to the lemons equilibrium. If these data are to be useful, we must be able to identify the source of the discrepancy between the data and the model's prediction.

Unfortunately, there are many potential reasons why the prediction was rejected:

1. The market did not reach an equilibrium. It is well-known in the experimental markets literature that equilibria are not immediately and instantaneously attained. Instead, it takes at least several market periods before the market "converges" to the predicted result. The number of periods necessary seems to vary with the complexity of the setting and this was not systematically studied. Smith (1976) found that simple double-auction markets reached a competitive equilibrium after two or three periods. In contrast, Forsythe and Lundholm (1990) found that as many as 30 periods were needed to reach a fully revealing rational expectations equilibrium in double-auction asset markets where agents had private information about an asset's true valuation.
2. The market reached a different equilibrium than that predicted by the lemons model. If subjects act as if they were in an infinite horizon game (since they did not know how many periods the experiment would last), then it is difficult to reject equilibrium behavior. For infinitely repeated games, the folk theorem (Fundenberg and Maskin 1986) states that any individually rational outcome can

be supported as a Nash equilibrium to the game. Thus, we would need to focus on a theory of equilibrium selection to reduce the set of candidate equilibria.

3. Buyers did not correctly process information about the quality of the service they received. Since buyers never observed the quality level directly but knew only whether a loss occurred, they may never have concluded that they were receiving low quality service and thus were willing to pay the higher prices associated with higher quality service. Better information about agent beliefs is needed to evaluate this alternative.
4. Buyers attached some credibility to a seller's disclosure of quality prior to sale even though sellers were under no obligation to provide this quality level. (An analysis of variance of the data provides some evidence on this.) In a single-period model such as the lemons model, the prediction is clear, disclosures should be completely disregarded by the buyers.[2] Did sellers' disclosures matter because of the multiperiod nature of the experiment or because buyers were fooled by seller puffery?
5. Perhaps the sellers' failure to always deliver low quality units was due to some moral stigma they felt about cheating the buyers. If this is the case, then either sellers' preferences were not adequately controlled in the experiment or they were not adequately modelled.[3] Yet again, the multiperiod nature of the experiment makes it impossible to test this explanation since these sellers may have been trying to establish good reputations in order to earn higher profits.
6. The higher than predicted prices for low quality units is due to a failure of the posted-offer auction that was conducted. This institution has been thoroughly tested when the items transacted have deterministic valuations to the buyers (see Ketcham et al. 1984). Although a posted-offer auction takes much longer to reach a competitive equilibrium than, say, a double-auction market, it nonetheless does so. However, to our knowledge, no one has examined the behavior of this auction when the items traded have stochastic valuations. Perhaps more than three sellers are needed in this case to get comparable performance from this institution.

We could expand the above list, but the point is that the data provide little help for theorists on how to modify existing theory. Consequently, the experiment is not a particularly useful one for improving our understanding of quality and seller disclosure in a market.

A Simpler Setting

In a single-period model, a buyer has no incentive to believe a seller's disclosures when the seller is under no obligation to tell the truth. Nonetheless, there was some limited evidence in DeJong et al. (1985) that buyers were often misled by sellers' dis-

---

[2] Further, sellers often offered to supply a particular quality of service for less than their cost. In a number of these instances, buyers accepted these offers even when lower offers were available to them. Would this have happened had buyers been explicitly told sellers' costs?

[3] Despite attempts to control subjects' preferences in our experiments, economists, like the psychologists before them, continue to observe behavior that demonstrates that some subjects are concerned with more than personal monetary reward. For example, the cooperation observed in prisoner dilemma games in psychology (Dawes 1980) continues to appear in economic experiments on public goods (Isaac et al. 1989) and other two-person games (Cooper et al. 1991).

Table 1
Subjects' Payoff Tables
Forsythe and Lundholm (1991)

Panel A. Seller's Payoffs:

	Low Quality		*Medium Quality*		*High Quality*	
	Sell	Not Sell	Sell	Not Sell	Sell	Not Sell
p_L	200	0	200	250	200	500
p_M	450	0	450	250	450	500
p_H	600	0	600	250	600	500

Panel B. Buyer's Payoffs:

	Low Quality		*Medium Quality*		*High Quality*	
	Sell	Not Sell	Sell	Not Sell	Sell	Not Sell
p_L	400	350	700	350	1000	350
p_M	150	350	450	350	750	350
p_H	0	350	300	350	600	350

closures (see point 4 above). We are unable to tell whether this reveals a fundamental flaw in the single-period model due to the design that was chosen.

Forsythe and Lundholm (1991) use a simpler design to address the issue of seller disclosure directly. Agent payoffs are induced in terms of utility by using the procedure first introduced by Roth and Malouf (1979) and further developed by Berg et al. (1986). Agents are randomly matched with one another from one period to the next. Prior to the start of a period, one agent in each pair is designated as the seller and the other as the buyer. Agents alternate from buyer to seller from one period to the next, and all agents know that the experimental session will last for 22 periods. Table 1 presents the seller and buyer payoffs. At the beginning of each period, each seller is endowed with one unit of a commodity that may be one of three quality types: low, medium, or high. The higher the quality type, the higher the unit's value to the buyer. Sellers incur no costs in connection with selling their units. After sellers learn the quality of the unit they have for sale, their paired buyers place a bid for the one unit by choosing a row. Rows represent the amount of the bid, which is one of three values, $p_L=200$, $p_M=450$, $p_H=600$. A seller then can either accept or reject the bid from the buyer by choosing a column. If the bid is accepted, the seller's payoffs do not depend on the unit's quality, but only on the price received. If the bid is rejected, the seller's payoffs increase with the asset's quality but do not depend upon the buyer's decision. Similarly, the buyer's payoffs are constant if the seller does not sell, regardless of the unit's quality. If the seller does sell, the buyer's payoffs decrease with the price and increase with the unit's quality. At the end of each period, agents are told their profits for the period and, if a purchase was made, the buyer is told the true quality of the unit purchased.

The treatment here is the restrictions on the disclosures that a seller can make prior to the buyer's entering a bid.[4] There are three levels of the treatment variable: The seller can make no disclosure, the seller can make any disclosure, and the seller can make a disclosure subject to an anti-fraud rule. An anti-fraud rule does not require the seller to disclose true quality but it does require that the true quality be included in any disclosure that is made.

There is a unique, sequential equilibrium for each of these treatments. With no disclosure, buyers should assume that the seller's unit is of the lowest quality. With no anti-fraud rule, the buyer should disregard the seller's disclosures and assume that the seller's unit is the lowest quality. In equilibrium, the buyer's beliefs are correct and the lemons prediction should result—in both treatments, only low quality units will be sold at low quality prices. With an anti-fraud rule, however, buyers should always assume that the true quality of a seller's unit is the same as the lowest quality disclosed (see Milgrom and Roberts 1986). In equilibrium, buyers' beliefs are correct, and all units are priced according to these beliefs. Further, the prices that buyers are willing to pay for each unit are sufficient to have sellers accept every bid, so all units are exchanged.

This design is simple. First, there is no confusion about whether a single-period or multiperiod model should apply. The design is a sequence of independent, one-shot games.[5] If we are going to test a single-period model such as the lemons model, it should be first tested as close as possible to its domain. Second, the environment (including preferences), institution, and set of agent action choices are all common knowledge. Third, there is a simple auction taking place for a single unit. The buyer's strategy set is limited to three prices. Fourth, buyers learn the actual quality of the seller's unit. Any failure of the model cannot be due to a buyer's failure to make correct inferences about the seller's true quality. Finally, there is no moral-hazard problem, there is only an adverse-selection problem. At this stage, one additional seller choice variable has been removed.

The research questions are also simplified and deal with the price and quality of units traded under each disclosure treatment, information content of seller disclosures with an anti-fraud rule, and whether voluntary unrestricted disclosures affect market performance. As the results in table 2 indicate, the model works quite well in predicting the observed outcomes when there is no disclosure and when there is an anti-fraud rule. That is, almost all buyers (92 percent) bid the low price when sellers cannot make a disclosure and, consequently, only low quality units are sold. When an anti-fraud rule is imposed on disclosure, almost all quality levels (96 percent) sell at the corresponding price (i.e., high quality units are sold at the high price, p_H; medium quality units at the medium price, p_M; and low quality units at the low price, p_L). Surprisingly, the data are much less supportive of the prediction that the seller's ability to make any disclosure without an anti-fraud rule has no effect on the adverse-selection problem. Buyers were still bidding the price associated with high quality items with some regularity even during the last five periods. Buyers bid more than the low quality price on 40 percent (20 out of 50) of all units and 43 percent (13 out of 30) of the units that the seller claimed to be high quality ones. Buyers are paying attention to what sellers say; they also bid

---

[4] Another treatment allows the seller to make offers that the buyer can accept. Since the nature of the single-period equilibria are more complicated in this latter case, we will discuss only the former.

[5] See Cooper et al. (1991) for a discussion of design issues and independent one-shot games.

Table 2
Prices Bid By Treatment
Forsythe and Lundholm (1991)

Panel A. No Disclosure Treatment—Distribution of Prices Bid—Last 5 Periods:

Prices Bid		
p_L	p_M	p_H
46	4	0

Panel B. Unrestricted Disclosure Treatment—Distribution of Prices Bid for Each Disclosure—Last 5 Periods:

Lowest Disclosed Quality	Prices Bid		
	p_L	p_M	p_H
Low	8	1	0
Medium	5	5	1
High	17	2	11

Panel C. Anti-Fraud Treatment—Distribution of Prices Bid for Each Disclosure—Last 5 Periods:

Lowest Disclosed Quality	Prices Bid		
	p_L	p_M	p_H
Low	27	1	0
Medium	1	12	0
High	0	0	9

higher prices than in the market with no disclosure. However, there is no significant relationship between the seller's disclosure and the quality of the unit.

Thus, in a very simple and well-controlled setting, one can demonstrate that buyers can be taken in by "cheap talking" sellers. This is particularly striking given that players alternate between being buyers and sellers each period. Nevertheless, this data provides evidence that pre-sales puffery can generate unwarranted optimism in buyers' beliefs. The design is also simple enough to explore potential explanations for this result and assist the theorist in modifying the model if necessary.

II. The Experimental Work in This Forum

The Forum articles examine issues of disclosure, auditing, and legal recourse. These are important issues in financial and audit markets where managers and auditors make disclosures about the quality of their product. While these issues are similar to those examined in section I, the research designs are very different, as is the complexity of the settings. As a consequence, when the results in both studies seem to contradict theory, it is not at all clear why theory fails.

Experimental Designs

The Wallin design consisted of a multiperiod setting in which a single seller and three buyers were matched for the duration of the experiment. At the beginning of each period, the reservation value of the asset to each buyer was randomly determined from a set of three possible values. This was the value of the asset if a buyer later purchased it and if it was a "winner." Next, sellers decided whether to make a costly investment that could improve the likelihood that the asset would be of high quality. A high quality asset had a higher likelihood of being a "winner" than a low quality asset. After making the investment decision and observing the asset's realized quality, the seller made a disclosure about the asset's true quality. Sellers could disclose the asset's quality as "high," "low," or "either," regardless of the truth. After learning the disclosure, the buyers placed bids for the asset in a second-price auction (i.e., one in which the highest bidder received the asset and paid the seller an amount equal to the second highest bid). Buyers then learned whether the asset was a winner. A winning asset paid the owner's reservation value and a losing asset was worthless. Each market lasted at least 25 periods, after which the decision to play an additional period was determined randomly. Although the model being tested is a single-period one, this design choice was an attempt to induce infinitely lived behavior. The treatments were the presence/absence of "robot" auditing and the presence/absence of legal recourse.

The Dopuch-King study is even more complex. Although redemption values are common to all buyers, there are now three economic agents involved in playing the same game for 20 periods. The new agent is the verifier (auditor). Each market has two sellers, two verifiers, and four buyers. The sequence of events and institutions are the same as in Wallin except that each period here began with a posted-offer auction in which verifiers announce offer prices to sellers, who then decide which verifier, if any, to hire. If hired, verifiers must decide whether to undertake a costly test of the seller's asset after its quality has been determined. The test determines asset quality without error. The verifier then makes a disclosure that either agrees or disagrees with the seller's quality disclosure. Verifiers could disagree with a seller's disclosure only if they had tested the asset.

The treatment variable in this study is the liability rule imposed on the verifier. There were three levels of this variable—no liability, negligence liability, and strict liability. The seller was never liable for his actions or disclosures. The predictions being tested are also from a single-period model.

These complex experiments create the types of concerns discussed earlier if the model performs poorly, and the six possible reasons listed in the preceding section for the failure of a similar model in the DeJong et al. study also apply here. Unfortunately, there was no attempt *ex ante* to control for these possible failures in the designs. In fact, the added features in these studies might make the occurrence of some of these reasons more likely. For example, a buyer's ability to make the correct inference about the seller's investment decision is further obscured by virtue of the fact that there are two lotteries—one determining quality and one determining whether the asset wins. As a second example, Dopuch and King add another strategic player and auction. The seller now must interact with two different sets of parties using two different trading mechanisms. In both auctions, the items traded have stochastic valuations—when deciding to accept or reject an auditor's offer, sellers presumably have some unmeasured prior beliefs about the likelihood that an auditor will test and buyers have some unmeasured prior beliefs on the likelihood that the asset they are bidding for is of high quality. Little to nothing is known about the performance of either the posted-offer auction (especially ones with only two sellers) or the second-price, sealed-bid auction for goods with stochastic valuations. The Wallin study uses only a single auction, but complicates it by allowing buyers to have different reservation values if the asset is a winner.

Analysis of Results

A closer look at the data in both studies reveals serious cohort effects.[6] These can be seen best in Wallin's tables 3–6. Examine cell A (no auditing/no legal recourse); even in late periods, the percentage of costly investment ranges from 18.2 to 91.7 percent. The percentage of truthful disclosure ranges from zero to 88.9 percent. Standard chi-square tests reject the ability to pool the observations across replications in each treatment on the basis of the costly investment decision and given that the asset was of low quality, the frequency of truthful disclosure in each cell except D (auditing/legal recourse available). As for the frequency with which an auditor is hired or a costly investment made in Dopuch and King, tests reject the ability to pool across replications in the last five periods of the three treatments. For example, in the strict liability treatment, the percentage of time the auditor is hired ranges from zero to 100 percent in these final periods.

The models in these two studies are not directly tested; only the qualitative effect of changing the level of a treatment variable is tested (e.g., the presence of auditing in Wallin or the negligence liability rule in Dopuch and King increases the frequency with which the seller makes the costly investment). Although we believe this to be a serious omission, the authors correctly argue that their single-period models would be rejected because neither model specifies an error distribution. Thus, a single observation contrary to, say, the hypothesis that the seller never undertakes the costly investment when both auditing and legal recourse (or the negligence and strict liability rule) are absent would lead us to reject the hypotheses from these models. Nonetheless, it is useful to know how much of the data a model explains. We would be less concerned about the efficacy of either model if the seller made the costly investment 5 percent of the time as opposed to 95 percent of the time.

In fact, the single-period models considered in these studies perform quite poorly. However, given the complexity of the research designs, it is hard to tell why this occurs. For example, do the models fail because the markets had not reached a single-period equilibrium? If they are in a single-period equilibrium, pooling observations over periods should be possible. However, for these two studies, each replication must be individually checked since replications differ within treatments. Using the chi-square test, we tested whether observations could be pooled over periods by replication. The pooling over periods done by Dopuch and King raises concerns for several of their replications. For the costly investment decision in the no liability treatment, tests reject the ability to pool over all periods in replication 2. In a similar vein, when the auditor hiring decision is considered, the last five periods of replication 2 differ from the previous periods, as does replication 4 in the strict liability treatment. Similar concerns exist for the Wallin data.[7] Since the models require equilibrium behavior, we are concerned that the data used from some of the replications are not equilibrium data.

---

[6] We thank the authors for making their data available to us.

[7] We are concerned by the arbitrary decision to pool all data from period 16 on. There are no tests reported to justify this pooling. However, our tests of the data indicate that the pooling is statistically justified after period 15 and could have been done from period 13 on if the author had so chosen.

Rather than use a setting wherein the model was given its best chance of "working," these experiments incorporated factors aimed at reducing the performance of the single-period models. If Wallin successfully induced behavior consistent with an infinite horizon environment, the differences observed across replications may be due to the different equilibria since numerous equilibria can exist in an infinite horizon environment. Similarly, in the multiperiod environment used in the study by Dopuch and King, different Nash equilibria can occur (apart from the subgame perfect one that they test).[8] One wonders whether these different equilibria can be used to explain the differences across replications or whether, alternatively, risk preferences differ across subjects. For example, perhaps subjects make systematic errors by overbidding for assets or overpricing audit services in the strict liability treatment because their risk preferences were not controlled. These are issues we wish the authors had explored.

The lack of empirical regularities undermines any policy conclusions that one might draw. Consider, for example replication B3 in the Wallin study where little auditing was purchased in late periods but the costly investment was always made and replication B2 where auditing was purchased two-thirds of the time in late periods but the costly investment was undertaken only in 50 percent of these cases. Which of these markets approximates the market of interest? Similarly, in the Dopuch and King study, is a market with a strict liability rule more like replication 2, where little or no costly investment is undertaken or auditing purchased, or more like replication 5, where a substantial amount of costly investment and auditing is undertaken? The presence of these cohort effects means one or more important variables were not controlled or accounted for.[9]

III. An Alternative Program of Research

The supply and demand for auditing are important issues in markets where managers make disclosures about product quality. The studies in this Forum offered a research plan that begins to address some of the issues in this area. In this section, we present a series of simple experimental designs to address these same issues. For convenience, these use the parameter set given in table 1 although this is not the only parameter set or design (nor necessarily the best) to use in all cases discussed below.

First, consider seller disclosure and the demand for auditing. When anti-fraud rules cannot be enforced without cost, there is a potential demand for auditing. For example, recall the results of Forsythe and Lundholm (1991). In the absence of anti-fraud rules, sellers' disclosures misled buyers and the market did not achieve the predicted equilibrium that only low quality units would sell at low quality prices.

To examine whether there is a demand for auditing here, the design could be expanded to introduce perfect auditing. For the moment, this abstracts from strategic issues in the supply of auditing, but it is a necessary omission if we are to determine whether the presence of asymmetric information alone is sufficient to lead sellers to purchase auditing that reveals quality perfectly.[10] To see what economic theory predicts here, return to table 1. If the seller must purchase auditing prior to observing quality, theory says that sellers must choose between two undominated alternatives. First, consider a seller who does not purchase auditing. As in the model outlined in section I, buyers will assume the worst and the seller will sell a unit only if it is a low quality one. The possible payoffs with this alternative are 500 if the unit is of high quality, 250 if it is of medium quality, and 200 if it is low. A seller's expected payoff from not purchasing auditing is 950/3(=(500+250+200)/3). Alternatively, a seller can purchase auditing, disclose the true quality level, and have this disclosure perfectly verified.[11] In this case, a seller will sell a unit and receive payoffs of 600 for a high quality unit, 450 if it is of medium quality and 200 if it is a low quality unit. The seller's expected payoff from purchasing auditing and disclosing truthfully is 1250/3(=(600 +450+200)/3). Thus, sellers should be willing to pay up to 100(=1250/3−950/3) to purchase auditing. If the cost is more, theory predicts that auditing will not be purchased.

---

[8] The work of Güth et al. (1982) and Ochs and Roth (1989) on ultimatum games provides strong experimental evidence against subgame perfection as an equilibrium concept.

[9] The presence of such cohort effects casts doubts on the replicability of the results with a different subject pool. It is not hard to imagine another researcher finding different results when the level of this unaccounted-for variable was different in this new subject pool.

[10] This is how auditing was first introduced in the theoretical literature, and its effect on such a setting is much easier to discern than when the setting is expanded to include strategic auditing. For example, see Evans (1980) and for a discussion of the literature Antle (1982).

This model generates a testable prediction about the demand for auditing which is nontrivial. Because of the payoffs sellers received in Forsythe and Lundholm (1991), sellers may continue to prefer settings in which they can disclose without having their disclosures verified. If buyers continue to be sufficiently gullible when reacting to sellers' unsubstantiated disclosures, there may be no incentives for sellers to purchase auditing. The profits that sellers earn from gullible buyers may exceed those they would earn with the purchase of perfect auditing.

Once any necessary modifications to the theory have been made in this initial setting with perfect auditing, one can then ask how the demand for auditing changes as the benefits from it are systematically varied. This can be done by examining different parameter sets from the one in table 1. Not all parameter sets lead to the theoretical prediction that only low quality units are sold in the absence of auditing. Also, the institution in which the asset is traded affects the benefits from auditing. So far, we have assumed that the buyer first bids for the asset and the seller then decides whether to sell. Alternatively, the seller could post a price that the buyer can accept. Given the parameters in table 1, the predicted equilibria between these two institutions differ in the absence of auditing.[12] When sellers post prices, some higher quality units are sold in equilibrium and so, in theory, the benefits from purchasing auditing are not as great.

Finally, the benefits from auditing depend on what the seller knows about the quality of the unit prior to purchasing auditing. This can affect the price of assets traded as well as which assets are traded. Suppose the seller knows the asset's quality prior to deciding to purchase auditing. A seller with a high quality asset would pay only up to 100 to purchase auditing since this is the difference in the payoffs from selling and not selling at the high price. A seller with a medium quality asset would pay up to 200 and a seller with a low quality asset would not pay anything to purchase auditing being assured of getting the low price. Thus, depending on the price structure of auditing, theory predicts that different sellers will purchase auditing, which in turn affects asset prices and trades between buyers and sellers. If the price for auditing does not exceed 100, a separating equilibrium is attained because sellers of high and medium quality assets will always purchase auditing, make truthful disclosures, and have the qualities verified. In equilibrium, all assets are traded, a buyer's belief about the quality of each purchased asset is correct, and all units are priced according to these beliefs. If the price of auditing is between 100 and 200, only sellers of medium quality assets will purchase auditing, make a truthful disclosure, and have the quality verified. The other sellers do not purchase auditing, both high and low quality are disclosed as high, the buyer offers the low price for each, and only the low quality asset is exchanged. As a consequence, gains from trading the high quality asset are not realized.

---

[11] If the seller purchases auditing and does not truthfully disclose the quality of the unit for sale, perfect auditing will result in a "disagree" report. In this case, the buyer should assume the worst and only bid the low price, p_l. Thus, a seller's payoffs will be the same as when auditing is not purchased. If the auditing is not free, then theory predicts that this strategy will never be observed since it is dominated by the seller not purchasing auditing.

[12] See Forsythe and Lundholm (1991) for an analysis of these two cases.

With a sufficient understanding of the demand for perfect auditing, one could next consider the supply aspects of auditing. This could be done by introducing a strategic auditor into the design. Hired auditors would decide whether to undertake a costly test of the asset's quality. Auditors who test report the true quality level of the seller's unit but auditors who do not test agree with the seller's disclosure. Analogous to the way the seller and buyer decide whether to exchange the asset and at what price, the auditor would begin by posting a price for the audit service. Auditors who test will report the true quality, but those who do not test must agree with the seller's disclosure. After receiving the disclosure and audit report, if one is available, the buyer proceeds to bid for the asset as before.

Clearly, the simple addition of a third strategic agent causes numerous complications. In the absence of any recourse against the auditor, the lemons model predicts that strategic auditors would never test in a one-shot game and that buyers would recognize the auditors' incentives and disregard the auditing report. In this case, auditing no longer makes sellers' disclosures credible. However, the data might reveal that gullible buyers disregard the auditors' disincentives to test and believe the audit report or that sellers also become gullible under these conditions if the presence of an auditor leads to a higher level of truthful disclosure than predicted.

All of the experiments proposed so far contain an adverse-selection problem, but moral hazard could be introduced if, for instance, one wanted to focus on disclosure and auditing when the seller makes a costly investment. Here, it would not be prudent to introduce this moral-hazard problem into the environment with perfect auditing. The simplest setting (i.e., one without auditing) would be appropriate initially to see the effect of disclosure and anti-fraud rules on a seller's decision in the presence of moral hazard. The introduction of auditing could follow later.

Although the experiments outlined here are much simpler than those of Wallin and Dopuch and King, it is nevertheless virtually certain that testing will reveal various "surprises." Nonetheless, a careful design will allow the cause of these surprises to be more readily identified so that the theory can be amended as necessary. Further, the proposed experiments have the potential to reveal much about the phenomena of interest-seller disclosure, the demand for auditing, and aspects of the supply of auditing with a strategic auditor.

IV. Summary and Conclusions

There should be a constant tension between theory and data, and experimental research allows testing of a theory as close to its domain as possible. A good experiment provides sufficient control to test a theory and offers insights about how the theory may be modified when the data contradict the theory's predictions. Because the purpose of much laboratory research is to create controlled conditions, it is not generally possible or desirable to re-create the "real world" in the laboratory. The complexity that mundane realism (Swieringa and Weick 1982) adds to a design impedes our understanding of the problems we want to study.

The alternative research agenda was designed to begin to study the issues chosen in this Forum. Since there is a well-developed theory to guide the research on these issues, we demonstrated that the methodology of experimental economics can be used to address this theory's predictions directly. It also has the potential to offer insights about where the theory performs poorly.

Our intent has been to convince the reader that the methods of experimental economics can offer new insights into the demand and supply of auditing. However, these methods must be used with the appropriate objectives and must be executed with sufficient control for the advantages they offer to lead to advances in the field of auditing.

References

Antle, R. 1982. The auditor as an economic agent. *Journal of Accounting Research* 20 (Autumn): 503–27.

Berg, J., D. Coursey, and J. Dickhaut. 1990. Experimental methods in accounting: A discussion of recurring issues. *Contemporary Accounting Research* 6 (Spring): 825–49.

———, L. Daley, J. Dickhaut, and J. O'Brien. 1986. Controlling preferences for lotteries on units of experimental exchange. *Quarterly Journal of Economics* 101 (May): 281–306.

Cooper, R., D. V. DeJong, R. Forsythe, and T. Ross. 1991. Cooperation without reputation. Working paper, University of Iowa, College of Business Administration.

Cox, J., and M. Isaac. 1986. Experimental economics and experimental psychology: Ever the twain shall meet? In *Economic Psychology: Intersections in Theory and Application*, edited by A. MacFadyen and H. MacFadyen. Amsterdam, the Netherlands: North-Holland.

Dawes, R. M. 1980. Social dilemmas. *Annual Review of Psychology* 31: 169–93.

DeJong, D. V., R. Forsythe, R. Lundholm, and W. Uecker. 1985. A laboratory investigation of the moral hazard problem in an agency relationship. *Journal of Accounting Research* 23 (Supplement): 81–120.

Evans, J. 1980. Optimal contracts with costly conditional auditing. *Journal of Accounting Research* 18 (Supplement): 108–28.

Forsythe, R., and R. Lundholm. 1990. Information aggregation in an experimental market. *Econometrica* 58 (May): 309–47.

———, and ———. 1991. Adverse selection and voluntary disclosure in financial markets: Some experimental evidence. Working paper, University of Iowa, College of Business Administration.

Fundenberg, P., and E. Maskin. 1986. The folk theorem in repeated games with discounting or with incomplete information. *Econometrica* 54 (May): 537–54.

Güth, W., R. Schmittberger, and B. Schwarze. 1982. An experimental analysis of ultimatum bargaining. *Journal of Economic Behavior and Organization* 3 (December): 367–88.

Isaac, R. M., J. M. Walker, and S. Thomas. 1989. Divergent evidence of free riding: An experimental examination of some possible explanations. *Public Choice* 43: 113–49.

Ketcham, J., V. L. Smith, and A. W. Williams. 1984. A comparison of posted-offer and double auction pricing institutions. *Review of Economic Studies* 51 (October): 595–614.

Milgrom, P., and J. Roberts. 1986. Relying on the information of interested parties. *Rand Journal of Economics* 17 (Spring): 18–32.

Ochs, J., and A. Roth. 1989. An experimental study of sequential bargaining. *American Economic Review* 79 (June): 355–84.

Plott, C. R. 1982. Industrial organization theory and experimental economics. *Journal of Economic Literature* 20 (December): 1485–527.

Roth, A., and M. W. K. Malouf. 1979. Game-theoretic models and the role of bargaining. *Psychological Review* 86 (November): 574–94.

Smith, V. L. 1976. Experimental economics: Induced value theory. *American Economic Review* 66 (May): 274–79.

———. 1982. Microeconomic systems as an experimental science. *American Economic Review* 72 (December): 923–55.

———, J. Schatzberg, and W. Waller. 1987. Experimental economics and auditing. *Auditing: A Journal of Practice & Theory* 7 (Fall): 71–93.

———, G. Suchanek, and A. Williams. 1988. Bubbles, crashes and endogenous expectations in experimental asset markets. *Econometrica* 56 (September): 1085–118.

Swieringa, R. J., and K. E. Weick. 1982. An assessment of laboratory experiments in accounting. *Journal of Accounting Research* 20 (Supplement): 56–101.

SHORTER ARTICLES

1. Jennifer J. Gaver, Kenneth M. Gaver, and George P. Battistel
   The Stock Market Reaction to Performance Plan Adoptions

2. Ashiq Ali, April Klein, and James Rosenfeld
   Analysts' Use of Information about Permanent and Transitory Earnings Components in Forecasting Annual EPS

3. Michael W. Maher, Peter Tiessen, Robert Colson, and Amy J. Broman
   Competition and Audit Fees

THE ACCOUNTING REVIEW
Vol. 67, No. 1
January 1992
pp. 172–182

The Stock Market Reaction to Performance Plan Adoptions

Jennifer J. Gaver
Kenneth M. Gaver
University of Georgia
George P. Battistel
University of Oregon

SYNOPSIS AND INTRODUCTION: This study examines the stock market reaction to the adoption of long-term compensation agreements for top management that are based on accounting goals. Prior researchers (Baril 1988; Larcker 1983; Smith and Watts 1982) argue that these agreements, known as performance plans, motivate executives to improve firm performance by working harder, lengthening their decision horizons, and becoming less risk-averse in their investment decisions. However, because the agreements are based on accounting goals, performance plans can distract managerial attention from the maximization of share price and can also encourage the manipulation of the accounting system (Gibbons and Murphy 1989). Managers may also reject projects with large, positive net present values in favor of less valuable projects with higher accounting profits.

In an earlier study, Larcker (1983) reports a significantly positive stock market reaction for a sample of 21 firms that adopted performance plans between 1971 and 1978. The reaction occurs on the trading day after the SEC receives the proxy statement announcing the plan adoption. However, we find no significant reaction during the two-day announcement period beginning with the SEC stamp date for a sample of 209 adoptions that occur between 1971 and 1980. Further, no evidence of abnormal stock performance is detected for the two-day periods beginning with either the date on which the board of directors voted to approve the plan or the proxy statement date.

We thank Gary Biddle, Pete Dukes, Susan Moyer, Eric Noreen, Terry Shevlin, D. Shores, and the participants of the UBCOW Accounting Conference for comments on an earlier draft. We also thank Larry Dann, Helen Gernon, Wayne Mikkelson, Megan Partch, Jim Peters, Peggy Wier, and Jerry Zimmerman for helpful suggestions. We acknowledge Lawrence Bickford, of Frederick W. Cook and Co., for generously providing compensation data. Finally, we thank A. Rashad Abdel-khalik, the editor, Richard Lambert, the associate editor, and two anonymous referees for many helpful comments and suggestions.

Submitted March 1990.
Accepted July 1991

Our study highlights the difficulties inherent in analyzing the stock market reaction to announcements made in proxy statements. First, pinpointing the timing of information dissemination is problematic. Second, firm-specific information unrelated to the event of interest is often released in the proxy statement and at the annual meeting. This complicates the interpretation of any unusual stock price behavior. In our study, significantly positive abnormal performance is observed for both adopting and nonadopting firms during the period beginning two days after the SEC stamp date and ending the day after the shareholders' meeting date. This suggests that the reaction observed for the adopting firms is due to the impending shareholders' meeting rather than to the performance plan adoption *per se*. We suggest that the results of any event study involving announcements made around the time of the annual shareholders' meeting should be interpreted with caution.

Key Words: *Accounting-based contracts, Executive compensation, Performance plans, Market-based event studies.*

Data Availability: *Data are available on request from the authors.*

THE remainder of the article is organized as follows. Section I describes the sample of performance plan adoptions analyzed in the study. Section II specifies the testing method and the announcement dates for the performance plan adoption. Section III presents the results of the study. Finally, section IV discusses the implications of the study for future research.

I. The Sample

A comprehensive list of 283 initial adoptions of performance plans between 1971 and 1980 was compiled from a search of the proxy statement holdings of the University of Oregon library, from the sample in Larcker (1983), and from information provided by Lawrence Bickford of Frederick W. Cook & Co., a compensation consulting firm. Five firms were dropped from the list because their proxy statements for the adoption year could not be obtained. An additional 40 firms were excluded because the required data were not available on the CRSP daily returns tape. This left 238 firms in the experimental sample.

More than 60 (*Value Line*) industries are represented in the 238-firm sample, with the largest concentration of adoptions in banking (10 firms), building (14 firms), food processing (18 firms), and petroleum (11 firms). Sample firms are generally larger than the general population of approximately 2,800 firms listed on the COMPUSTAT tapes. For 83.2 percent of the sample, total assets measured in the year prior to adoption are above the median asset size for all COMPUSTAT firms measured in the same year. The sample firms are also more profitable than the COMPUSTAT firms in general. For 65.1 percent of the sample, return on assets in the year prior to adoption is above the median for all COMPUSTAT firms measured in the same year. However, financial leverage and liquidity of the sample firms are similar to those of other COMPUSTAT firms.

Also, in the year prior to adoption, all sample firms paid a fixed salary to their executives, and the overwhelming majority also made annual bonus awards and stock option grants.[1]

II. Experimental Design

Testing Method

Abnormal stock price performance associated with the performance plan adoption is examined by using standard market model methodology (Patell 1976). Specifically, it is assumed that in the absence of price effects related to plan adoption, security returns are characterized by the single factor market model:

$$R_{jt}=\alpha_j+\beta_j R_{mt}+\epsilon_{jt},$$

where R_{jt} is the rate of return for firm j on day t, R_{mt} is the market return on day t, β_j is the systematic risk of firm j, α_j is the rate of return on firm j when R_{mt} is zero, and ϵ_{jt} is a serially independent disturbance term with $E(\epsilon_{jt})=0$. The market model parameters for each firm, α_j and β_j, are estimated over the period from day -170 to day -21 relative to each possible announcement date. The CRSP daily returns file and the CRSP equally weighted market index are used for returns data.

The abnormal daily return, e_{jt}, for each firm during each day in the announcement period is estimated as:

$$e_{jt}=R_{jt}-a_j-b_j R_{mt},$$

where a_j and b_j are the least squares estimates of α_j and β_j. Abnormal returns for the firm j on day t are standardized by the estimated standard deviation of the residuals from the market model and a factor reflecting the increase in variability due to predicting outside of the estimation period. These standardized abnormal returns are then averaged cross-sectionally and multiplied by the square root of the number of firms to produce a z-statistic that is used to test the hypothesis of no abnormal performance during day t in the announcement period (see Patell 1976).

The significance of the cumulative reaction during the event period is tested in an analogous fashion. Abnormal returns are first accumulated over the announcement period and then divided by the standard error to produce a standardized, cumulative abnormal return (SCAR) for each firm. These are then averaged cross-sectionally and multiplied by the square root of the number of firms to produce a z-statistic that is used to test the hypothesis of no abnormal performance during the announcement period.[2]

Announcement Dates

Because the timing of the release of information about the performance plan adoption is uncertain, the market reaction is examined during three alternative event periods consisting of an event date and the following trading day. The three event dates

---

[1] The results of Gaver (forthcoming) and Lambert and Larcker (1987) suggest that firms adopt performance plans when they reach a mature phase in their life cycles. For most firms, the adoption occurs after a stock option plan is in place, and we found little evidence that performance plans replace stock option plans. This suggests that the performance plan adoption will not significantly extend the manager's planning horizon.

[2] This procedure for standardizing the cumulative abnormal returns is described by Mikkelson and Partch (1988) and attributed to Craig Ansley.

(in chronological order) are the board of directors' meeting date, the proxy statement date, and the SEC stamp date.[3] The board meeting date, which is the date that the board of directors voted to approve the performance plan, is disclosed by 110 firms in the sample. The proxy statement date is the date that the proxy statement is signed. Typically, this is also the date that the statement is mailed to stockholders. This date is disclosed by all of the 238 sample firms. The SEC stamp date is the date that the proxy statement is stamped as received by the SEC. A readable SEC stamp date is available for 209 of the firms.[4]

A disadvantage of using two-day announcement intervals is that they may be too short to capture the stock price response to the announcement of the performance plan adoption. In their study of changes in long-term compensation contracts, Brickley et al. (1985) address this issue by analyzing stock price behavior over two longer intervals: from the board meeting date through the day after the SEC stamp date and from the second day after the SEC stamp date through the day after the shareholders' meeting date.[5] We report the results of a similar analysis in this study. Of the 238 sample firms, 98 disclose the board meeting date and also have readable SEC stamp dates. The SEC stamp date and the shareholders' meeting date are both available for 209 firms.

III. Results

Table 1 presents the results for the three two-day announcement periods. The first column indicates the event date and the size of each subsample, and is followed by the average abnormal performance for each day in the interval, as well as the average two-day cumulative abnormal return (CAR). The third column reports the z-statistic for testing the null hypothesis of no abnormal performance during the event period, and the significance level of the z-statistic is given in the fourth column. The last column reports the significance level of the Wilcoxon signed ranks test. No significant price reaction is detected around the board meeting date, the proxy statement date, or the SEC stamp date (p-values associated with the z-statistics range from 0.154 to 0.657).

One explanation for the insignificant results is that, for many firms in the sample, the adoption of a performance plan is expected by market participants. As a result, the anticipated benefits are impounded in stock price prior to the proxy release date. To investigate this possibility, the 238-firm sample is partitioned according to the degree to which adoption is likely to be anticipated by market participants.

Three subsamples are examined: the early adopters, the industry pacesetters, and the intersection of the early adopters among the industry pacesetters. A review of the

---

[3] In their study of long-term incentive plans, Brickley et al. (1985) also examine the two-day reaction on the shareholders' meeting date. We do not use this date as a possible event date for two reasons. First, the meeting date follows the proxy statement date by a median of 25 trading days in our sample. Any proxy statement disclosures would already be impounded in security prices well before the meeting date. Second, Brickley (1986) reports significantly positive abnormal stock performance for a random sample of firms on their shareholders' meeting date. Thus, any response detected on that date would be difficult to interpret.

[4] For the sample, the median number of trading days between the board meeting date and the proxy statement date is 55 (mean 65, range: 5 to 280). The median number of trading days between the proxy statement date and the SEC stamp date is two (mean 2.5, range: −9 to 18). Finally, the median number of trading days between the SEC stamp date and the shareholders' meeting date is 22 (mean 24.4, range: 6 to 57).

[5] Although, as we argue, the market will impound the news of the performance plan adoption into stock price prior to the stockholders' meeting date, it is not clear exactly when this will occur. We extend our examination to the day following the shareholders' meeting date to maintain consistency with Brickley et al. (1985).

Table 1
Abnormal Return During Two-Day Announcement Periods Associated with the Board Meeting Date, the Proxy Statement Date, and the SEC Stamp Date[a]

Event Date	*Abnormal Return (Percentage)*	*z-statistic*	*p-value, z-statistic*	*p-value, Wilcoxon*[b]
Board Meeting	−0.2	−0.821	0.412	0.524
Board+1 day	−0.1	−0.711	0.477	0.380
Two-day CAR[c] (n=110)	−0.3	−1.072	0.284	0.084
Proxy Statement[d]	0.2	1.426	0.154	0.296
Proxy+1 day	−0.1	−0.783	0.434	0.404
Two-day CAR (n=238)	0.0	0.444	0.657	0.590
SEC Stamp[e]	0.0	0.579	0.562	0.821
SEC+1 day	0.0	0.480	0.631	0.694
Two-day CAR (n=209)	0.0	0.744	0.457	0.799

[a] The full sample consists of 238 firms that adopted performance plans between 1971 and 1980, have proxy statements available for the adoption year, and have data on the CRSP daily returns tape. Of these firms, 110 disclose the date on which the board of directors voted to approve the performance plan, and 209 have readable SEC stamp dates.

[b] The significance level of the Wilcoxon signed ranks test.

[c] CAR=cumulative abnormal return.

[d] The median number of trading days between the board meeting date and the proxy statement date is 55 (mean 55, range: 5 to 280).

[e] The median number of trading days between the proxy statement date and the SEC stamp date is two (mean 2.5, range: −9 to 18). The median number of trading days between the SEC stamp date and the shareholders' meeting date is 22 (mean 24.4, range: 6 to 57).

sample reveals that 127 of the 238 adoptions are clustered in 1978–80, the last three years of the sample period. Thus, we define the early adopters as those firms that adopt performance plans between 1971 and 1977. This subsample consists of 111 firms. Next, we define the industry pacesetters as those firms that are the first in their (three-digit SIC code) industry to adopt a performance plan. Coincidentally, this classification also produces a subsample of 111 firms. Finally, we examine the intersection of the early adopters among the industry pacesetters. This subsample is comprised of 72 firms.

Table 2 presents abnormal performance during the two-day periods beginning with the board of directors' meeting date, the proxy statement date, and the SEC stamp date for these subsamples. No significant market reaction is observed for any subsample around either the proxy statement date or the SEC stamp date. However, both the z-statistic and the Wilcoxon signed ranks test indicate that a significantly *negative* reaction occurs for the early adopters during the two-day period that begins with the board meeting date (the z-statistic for the two-day CAR is −3.055; *p*-value for the Wilcoxon is 0.012).

It is possible that the negative reaction at the board meeting date is due to other nonroutine proxy statement agenda items that accompany performance plan

Table 2
Abnormal Return During Two-Day Announcement Periods for Subsamples of Adoptions Most Likely Unanticipated by the Market[a]

Event Date	*Abnormal Return (Percentage)*	*z-statistic*	*p-value, z-statistic*	*p-value, Wilcoxon*[b]
Early Adopters				
Board Meeting	−0.5	−1.874	0.061	0.324
Board+1 day	−0.5	−2.479	0.013	0.022
Two-day CAR[c]	−1.0	−3.055	0.002	0.012
(n=56)				
Proxy Statement	0.3	1.368	0.171	0.249
Proxy+1 day	−0.2	−0.564	0.573	0.429
Two-day CAR	0.1	0.565	0.572	0.609
(n=111)				
SEC Stamp	−0.2	−0.688	0.492	0.289
SEC+1 day	0.2	1.206	0.228	0.305
Two-day CAR	0.1	0.369	0.712	0.977
(n=104)				
Pacesetters				
Board Meeting	−0.4	−1.742	0.081	0.260
Board+1 day	−0.2	−1.121	0.262	0.432
Two-day CAR	−0.6	−2.014	0.044	0.062
(n=58)				
Proxy Statement	−0.1	−0.225	0.822	0.832
Proxy+1 day	−0.2	−0.074	0.941	0.663
Two-day CAR	−0.3	−0.202	0.840	0.789
(n=111)				
SEC Stamp	−0.2	−0.499	0.618	0.307
SEC+1 day	0.1	0.195	0.845	0.885
Two-day CAR	−0.2	−0.216	0.829	0.423
(n=100)				

(continued)

adoptions.[6] To investigate this possibility, we partition the sample of 56 early adoptions with available board meeting dates into two subsamples: the 25 firms whose performance plan adoption is the only nonroutine item on the proxy statement agenda, and the 31 remaining firms whose performance plan adoption is accompanied by other nonroutine agenda items, including proposals to amend the corporate charter, to approve acquisitions, to change the voting procedure for the election of board members, and to make additional modifications in executive compensation agreements.

---

[6] We define routine proxy statement agenda items as the election of board members, the ratification of auditors, and miscellaneous shareholder proposals. These items are routinely included on the agenda of annual meetings of virtually all firms.

Table 2—*Continued*

Abnormal Return During Two-Day Announcement Periods for Subsamples of Adoptions Most Likely Unanticipated by the Market[a]

Event Date	Abnormal Return (Percentage)	z-statistic	p-value, z-statistic	p-value, Wilcoxon[b]
Early Adopters and Pacesetters				
Board Meeting	−0.4	−1.455	0.146	0.528
Board+1 day	−0.4	−1.651	0.099	0.232
Two-day CAR[c] ($n=38$)	−0.8	−2.184	0.029	0.078
Proxy Statement	0.1	0.820	0.412	0.582
Proxy+1 day	−0.3	−0.687	0.492	0.426
Two-day CAR ($n=72$)	−0.2	−0.095	0.924	0.920
SEC Stamp	−0.3	−0.890	0.374	0.231
SEC+1 day	0.3	1.133	0.257	0.366
Two-day CAR ($n=63$)	0.0	0.176	0.861	0.924

[a] The full sample consists of 238 firms that adopted performance plans between 1971 and 1980, have proxy statements available for the adoption year, and have data on the CRSP daily returns tape. Of these firms, 111 adopted performance plans prior to 1978. Coincidentally, 111 firms were the first in their (three-digit SIC code) industry to adopt a plan. Seventy-two firms were the first in their industry to adopt and also adopted prior to 1978.

[b] The significance level of the Wilcoxon signed ranks test.

[c] CAR = cumulative abnormal return.

Table 3 presents abnormal performance during the two-day period beginning with the board meeting date for these subsamples. Panel A reports the results for the subsample of 25 firms. No significant reaction is observed around the board meeting date for this subsample (the z-statistic for the two-day CAR is −0.924). Panel B reports the results for the remaining set of 31 firms, and a significantly negative reaction is observed for this subsample (the z-statistic for the two-day CAR is −3.277). These results suggest that the significantly negative reaction at the board meeting date reported in table 2 is not due to the performance plan adoption *per se*, but appears to be associated with other proxy statement agenda items.

Table 4 presents the CARs obtained when using the longer event periods investigated by Brickley et al. (1985). Panel A reports results for the period beginning with the board meeting date and ending the day after the SEC stamp date for the subsample of 98 adoptions with both dates available. No significant reaction is observed during this interval (z-statistic = −0.734). In contrast, a significantly positive reaction is found during the period beginning two days after the SEC stamp date and ending the day after the shareholders' meeting date for the subsample of 209 adopters with SEC stamp dates available (z-statistic = 3.476).[7]

---

[7] Our results differ from those of Brickley et al. (1985), who find a significantly positive reaction in the interval between the board meeting and SEC stamp dates for their sample of changes in long-term incentive plans. However, Brickley et al. use a procedure for standardizing cumulative abnormal returns that ignores the serial

Table 3
Abnormal Return During Two-Day Announcement Periods Associated with the Board Meeting Date for Subsamples of Pre–1978 Adopters[a]

Panel A. Firms Whose Plan Adoption is the Only Nonroutine Proxy Statement Agenda Item (n=25):

Event Date	*Abnormal Return (Percentage)*	*z-statistic*	*p-value, z-statistic*	*p-value, Wilcoxon*[b]
Board Meeting	−0.2	−0.668	0.504	0.600
Board+1 day	−0.1	−0.645	0.519	0.427
Two-day CAR[c]	−0.3	−0.924	0.355	0.242

Panel B. Firms Whose Plan Adoption is Accompanied by Other Nonstandard Proxy Statement Agenda Items (n=31):

Event Date	*Abnormal Return (Percentage)*	*z-statistic*	*p-value, z-statistic*	*p-value, Wilcoxon*
Board Meeting	−0.6	−1.919	0.055	0.357
Board+1 day	−0.8	−2.753	0.006	0.013
Two-day CAR	−1.5	−3.277	0.001	0.015

[a] The full sample consists of 238 firms that adopted performance plans between 1971 and 1980, have proxy statements available for the adoption year, and have data on the CRSP daily returns tape. A total of 56 sample firms adopted performance plans between 1971 and 1977 (that is, pre–1978) and also disclose board meeting dates. In 25 cases, the adoption year proxy statement does not contain any nonroutine agenda items other than the performance plan adoption; the remaining 31 adoptions are accompanied by other nonstandard proposals. We define nonroutine agenda items as anything other than (1) the election of board members, (2) the ratification of auditors, and (3) shareholder proposals.

[b] The significance level of the Wilcoxon signed rank test.

[c] CAR=cumulative abnormal return.

To further investigate the significant positive reaction during the period leading up to the shareholders' meeting date, we select a control firm for each firm in our sample. Control firms are in the same two-digit SIC classification as the related sample firms, fall within two deciles of asset size relative to the population of COMPUSTAT firms, but have not adopted a performance plan.[8] Table 5 presents a comparison of the CARs for the experimental and control samples during the period beginning two days after the SEC stamp date and ending the day after the shareholders' meeting. The analysis is based on 185 matched pairs, with returns aligned in event time.[9] We observe highly significant, positive CARs for both the experimental firms (z-statistic, 3.547) and the control firms (z-statistic, 3.553) during this interval. Paired comparison tests indicate no significant difference between the returns behavior of the two groups. This comparison

---

dependencies that arise from using a common estimation model to obtain the daily abnormal returns. As a result, their test statistic is biased toward rejection of the null hypothesis of no abnormal performance. The correct method for standardizing cumulative abnormal returns, which takes these dependencies into account, is described by Mikkelson and Partch (1988). We use this method in our tests.

[8] In most cases, the control firm is actually in the same three- or even four-digit SIC industry code as the matched adopter.

[9] Although SEC stamp dates and shareholders' meeting dates are available for 209 experimental firms, the analysis is conducted with 185 matched pairs because of missing SEC stamp dates for some control firms.

Table 4

Cumulative Abnormal Return (CAR) During Longer Event Periods for Adoptions with Known Board Meeting and SEC Stamp Dates

Panel A. Trading Interval Beginning with Board Meeting Date and Ending the Day After the SEC Stamp Date (n=98):

Average Interval[a]	*Average CAR*	*Average Standardized CAR*	*z-statistic*[b]	*% Positive*[c]	*p-value, Wilcoxon*[d]
60.265	−0.008	−0.074	−0.734	49.0	0.493

Panel B. Trading Interval Beginning Two Days After SEC Stamp Date and Ending Day After Shareholders' Meeting (n=209):

Average Interval	*Average CAR*	*Average Standardized CAR*	*z-statistic*	*% Positive*	*p-value, Wilcoxon*
22.5	0.023	0.240	3.476[e]	57.9[e]	0.001

Note: The full sample consists of 238 firms that adopted performance plans between 1971 and 1980, have proxy statements available for the adoption year, and have data on the CRSP daily returns tape. Of the 238 firms, 209 have readable SEC stamp dates. Of these 209 firms, 98 also disclose the date on which the board of directors voted to approve the performance plan adoption.

[a] Average number of days in the interval. Note that the length of this interval varies across firms.

[b] The z-statistic for the average standardized cumulative abnormal return for the sample.

[c] The percentage of the sample with a positive cumulative abnormal return for the interval.

[d] The significance level of the Wilcoxon signed ranks test.

[e] Significant at the 0.01 level, two-tailed test.

suggests that the significant positive reaction for the experimental firms is associated with the upcoming shareholders' meeting rather than with the performance plan adoption *per se*.

IV. Conclusion

Our results are inconsistent with the prior conclusion in the accounting literature that performance plan adoptions significantly increase shareholder wealth. One explanation for the findings is that the adoption is viewed by market participants as a relatively minor change in the compensation contract that will not significantly alter managerial incentives. This explanation seems plausible since most adopting firms have pre-adoption contracts such as bonus and stock option plans, which arguably reduce the marginal incentive benefits of the performance plan. Also, internal employment contracts are only one means by which potential conflicts between owners and managers are allayed. Alternative monitoring mechanisms include oversight of managers by the board of directors and the disciplining forces of competition in managerial labor markets (Fama 1980), the firm's product markets (Hart 1983), and the market for corporate control (Manne 1965). Further, nonpecuniary rewards associated with successfully managing the firm also are likely to motivate managers to act in the shareholders' interests. These forces tend to mitigate any stock price response to performance plan adoption.

Table 5
Comparison of Cumulative Abnormal Return (CAR) for Experimental and Control Samples[a] During the Period Beginning Two Days after the SEC Stamp Date and Ending the Day after the Shareholders' Meeting Date

Sample	*Average Interval*[a]	*Average CAR*	*Average Standardized CAR*	*z-statistic*[b]	*% Positive*[c]	*p-value, Wilcoxon*[d]
Experimental	22.5	0.025	0.261	3.547*	61.6*	0.000
Control	20.4	0.022	0.261	3.553*	57.8*	0.003

Paired Comparison Tests: Average difference in standardized CAR, 0; t-statistic, 0; median difference in standardized CAR, 0.085; and z-statistic, 0.470. Differences are calculated by subtracting the standardized cumulative abnormal return of the control firm from the standardized cumulative abnormal return of the experimental firm.

Note: Experimental firms adopted performance plans between 1971 and 1980 and have readable SEC stamp dates on their adoption year proxy statements. A control firm is selected for each experimental firm on the basis of asset size and industry membership. Although the experimental sample consists of 209 firms, the analysis is conducted on 185 matched pairs because of missing SEC stamp dates for some control firms. Returns for experimental and control firms are aligned in event time.

[a] The average number of trading days in the interval for each sample. Note that the length of the interval varies across firms.

[b] The z-statistic for the average standardized cumulative abnormal return for the sample.

[c] The percentage of the sample with a positive cumulative abnormal return for the interval.

[d] The significance level of the Wilcoxon signed rank test.

* Significant at the 0.01 level, two-tailed test.

It is possible that, unlike performance plans, initial adoptions of stock option plans or bonus plans have discernibly positive shareholder-wealth effects. Initial evidence that bonus plan adoptions are associated with significantly positive abnormal price performance is provided by Tehranian and Waegelein (1985). However, their study is based on a small sample (42 firms) and uses monthly returns data. Further, the results of our study indicate that the upcoming shareholders' meeting can be a seriously confounding event in studies that attempt to measure the stock market response to disclosures made in the annual proxy statement. The results that we observe seem to be consistent with the "annual meeting effect" first observed by Brickley (1986). Attributing positive stock price reactions to announcements of compensation plan adoptions must therefore be approached with caution.

References

Baril, C. P. 1988. Long-term incentive compensation, ownership, and the decision horizon problem. Working paper, McIntire School of Commerce, University of Virginia, Charlottesville.

Brickley, J. A. 1986. Interpreting common stock returns around proxy statement disclosures and annual shareholder meetings. *Journal of Financial and Quantitative Analysis* 21 (September): 343–49.

———, S. Bhagat, and R. C. Lease. 1985. The impact of long-range managerial compensation plans on shareholder wealth. *Journal of Accounting & Economics* 7 (April): 115–29.

Fama, E. F. 1980. Agency problems and the theory of the firm. *Journal of Political Economy* 88 (April): 288–307.

Gaver, J. J. 1991. Incentive effects and managerial compensation contracts: A study of performance plan adoptions. *Journal of Accounting, Auditing and Finance* (forthcoming).

Gibbons, R., and K. J. Murphy. 1989. Agency theory and the determinants of executive compensation. Working paper, University of Rochester, NY.

Hart, O. D. 1983. The market mechanism as an incentive scheme. *Bell Journal of Economics* 14 (Autumn): 366–82.

Lambert, R. A., and D. F. Larcker. 1987. An analysis of the use of accounting and market measures of performance in executive compensation contracts. *Journal of Accounting Research* 25 (Supplement): 85–125.

Larcker, D. F. 1983. The association between performance plan adoption and corporate capital investment. *Journal of Accounting & Economics* 5 (April): 9–30.

Manne, H. G. 1965. Mergers and the market for corporate control. *Journal of Political Economy* 73 (February): 110–20.

Mikkelson, W. H., and M. M. Partch. 1988. Withdrawn security offerings. *Journal of Financial and Quantitative Analysis* 23 (June): 119–34.

Patell, J. M. 1976. Corporate forecasts of earnings per share and stock price behavior: Empirical tests. *Journal of Accounting Research* 14 (Autumn): 246–76.

Smith, C. W., and R. L. Watts. 1982. Incentive and tax effects of U.S. executive compensation plans. *Australian Journal of Management* 7 (December): 39–157.

Tehranian, H., and J. F. Waegelein. 1985. Market reaction to short-term executive compensation plan adoption. *Journal of Accounting & Economics* 7 (April): 131–44.

THE ACCOUNTING REVIEW
Vol. 67, No. 1
January 1992
pp. 183–198

Analysts' Use of Information about Permanent and Transitory Earnings Components in Forecasting Annual EPS

Ashiq Ali
Columbia University and New York University
April Klein
New York University
James Rosenfeld
Emory University

SYNOPSIS AND INTRODUCTION: An intriguing anomaly in recent market-based accounting research is that both the market and analysts appear not to recognize properly the time-series properties of quarterly earnings shocks. Bernard and Thomas (1990) and Freeman and Tse (1989) present evidence that the market underestimates the implications of previous period earnings for future earnings. Mendenhall (1991) and Abarbanell and Bernard (1991) find that analysts do not seem to utilize time-series information about earnings correctly when setting their forecasts. Specifically, these last two studies document a positive serial correlation in analysts' quarterly forecast errors and interpret this finding as analysts systematically underestimating the persistence of past earnings forecast errors in forecasting future earnings.

The purpose of this article is to examine whether analysts properly recognize the time-series properties of *annual* earnings when setting their estimates of future earnings. Givoly (1985) investigates this issue and finds that analysts' forecasts of annual earnings per share (EPS) are unbiased and that prediction errors are not serially correlated. He concludes that

We gratefully acknowledge the valuable comments and suggestions from George Benston, Larry Brown, Richard Mendenhall, Shehzad Mian, Jacob Thomas, Nicholas Valerio, Larry Wall, Greg Waymire, two anonymous referees, and the participants of the NYU Accounting and Emory Finance Workshops. We also would like to thank IBES for supplying us with the analysts' forecast data.

Submitted January 1990.
Accepted July 1991.

forecasts are formed in an efficient manner. In contrast, this study finds that, on average, analysts set overly optimistic estimates of the next period's annual EPS and that forecast errors display significantly positive serial correlation. These results hold for short-term as well as for longer-term IBES consensus forecasts.

Bias and positive serial correlation in forecast errors suggest that analysts do not properly recognize the time-series properties of earnings when setting expectations of future earnings. Studies show that, for any given year, earnings shocks have both permanent and transitory (i.e., mean-reverting) components (see, e.g., Brooks and Buckmaster 1976; Ou 1990; Ou and Penman 1989b) and that the level of earnings persistence varies across firm-years. We examine whether analysts are aware of the differences between permanent and temporary components in the previous year's earnings when predicting future earnings.

The findings show that analysts are able to differentiate partially between permanent and temporary components in previous period earnings. We also find that the overestimation bias in forecasts is most pronounced for firms that recently experienced negative earnings. In contrast, the positive serial correlation is most evident for firms that had predominantly permanent earnings. These results suggest that the overestimation bias and serial correlation are not uniform across firms.

To address the economic significance of our findings, estimates of bias and serial correlation in forecast errors are used to adjust existing forecasts. The accuracy of the adjusted forecasts are then compared to the unadjusted forecasts. A significant improvement in forecasting ability is found for the adjusted longer-term forecasts, as evidenced by a 12 percent improvement in the mean squared error. Further, the bias and serial correlation in the adjusted forecasts are significantly less than in the actual forecasts for both longer and short terms. These results further support the view that analysts do not utilize available information efficiently when setting forecasts.

Key Words: *Analysts' forecast errors, Overprediction bias, Serial correlation in forecast errors, Permanent and temporary components of earnings.*

Data Availability: *Analysts' forecast data are available from Lynch, Jones and Ryan's* Institutional Brokers Estimate System (IBES) *1990 database. Stock market prices and returns are available on the* Center for Research in Security Prices (CRSP) *monthly and CRSP NASDAQ databases.*

THE article is organized as follows. Section I describes the data and the test variables. Section II documents the overprediction bias and serial correlation in longer-term forecast errors for the entire sample. Section III examines whether analysts utilize the time-series properties of earnings when setting forecasts. Section IV explores the economic implications of our findings. Section V repeats the analyses with one-month forecasts, and section VI presents our conclusions.

I. Data and Variable Definitions

Analysts' forecasts errors for firm i for year t, $FEPR_{it}$, is defined as:

$$FEPR_{it}=(EPS_{it}-F_{it})/P_{it}, \tag{1}$$

where EPS_{it} is reported annual earnings per share for firm i for year t, F_{it} is the median analysts' forecast for annual EPS of year t, and P_{it} is the market price of the stock at the beginning of the month in which analysts' forecasts are released by IBES.[1] The initial sample of analysts' forecasts consists of firms included in the IBES 1990 database, which contains forecasts of annual earnings from 1976 to 1990. We use the median forecast as our measure of the analysts' consensus forecast. To help insure that the median is a meaningful statistic, we require each firm to have at least three earnings estimates. Short-term forecasts for year t are taken one month prior to the end of fiscal-year t. Longer-term forecasts are taken eight months prior to the end of the fiscal year. Selection of an eight-month forecast horizon is based on our requirement that the previous year's EPS be known by analysts when making their current year's forecast.[2] We further require that EPS for year $t-1$ be known at the time the EPS forecast for year t is made.

Stock prices are taken from the *Center for Research in Security Prices* (CRSP) monthly database (which contains all firms listed on the NYSE) and the CRSP NASDAQ database (which contains over-the-counter firms through 1988). If prices are unavailable from either source, the IBES database is used. (The IBES database does not contain price information before 1983.)[3]

To be included in the final sample, each firm must have at least three consecutive years of data (from year $t-2$ to year t) for reported and forecasted EPS, as well as price data at the beginning of the forecast month. We do not limit our sample on the basis of listing status and/or fiscal year-end month.[4] Retaining OTC firms allows for the inclusion of smaller firms; including firms with non-December fiscal year-ends produces a sample with a broader number of industries (Smith and Pourciau 1988).

These restrictions yield a sample of 5,365 firm-years over the 12-year period 1978–1989. Table 1 presents a breakdown of the firm-years according to firms' listing status as well as the month of their fiscal year-ends.

For the sample, the mean and median eight-month forecast errors (deflated by price) are -3.02 and -0.24, significant at the 0.01 level, which suggest that analysts' forecasts are, on average, upwardly biased over the period 1978–1989.[5] The Pearson and Spearman correlation coefficients between $FEPR_{it}$ and $FEPR_{it-1}$ are 0.36 and 0.33,

---

[1] To evaluate the sensitivity of our analyses to the selection of the deflator, we re-estimated our equations with the standard deviation of analysts' forecasts at the time of the forecast (available from IBES) as an alternate deflator and also on raw, undeflated data. The results with these two measures are qualitatively the same as those reported in this article and can be obtained from the authors.

[2] Penman (1987) reports that 92 percent of all annual reports are filed with the SEC within 12 weeks after the end of the fiscal year. We allow an extra month in view of O'Brien's (1988) finding that an average lag of 34 trading days exists between the analysts' forecast date and the date of the forecast's publication in the IBES database.

[3] All of the data are adjusted for stock splits and stock dividends.

[4] Givoly (1985) restricts his sample to 27 firms listed on the NYSE with December 31 fiscal year-ends.

[5] The eight-month forecast error has a maximum value of 102.25 and a minimum value of -631.13. The first and third quartiles of the distribution are -2.08 and 0.33, respectively.

Table 1
Distribution of Sample Observations According to Firms' Listing Status and Month of Fiscal Year-End (1978–1989)

Firm Characteristics	*Firm-Years*	*Percentage Firm-Years*
Listing Status		
New York Stock Exchange	4,195	78.2
American Stock Exchange	188	3.5
Over-the-Counter	982	18.3
Total	5,365	100.0
Fiscal Year-End		
January	134	2.5
February	59	1.1
March	129	2.4
April	91	1.7
May	80	1.5
June	381	7.1
July	97	1.8
August	96	1.8
September	322	6.0
October	150	2.8
November	86	1.6
December	3,740	69.7
Total	5,365	100.0

significant at the 0.01 level, and are preliminary evidence that longer-term annual forecast errors may be serially correlated.[6]

The mean and median one-month forecast errors are −1.38 and −0.03, each significant at the 0.01 level.[7] The Pearson and Spearman correlation coefficients between yearly forecast errors are 0.27 and 0.18, also significant at the 0.01 level. Thus, both long-term and short-term forecast errors display evidence of bias and first-order serial correlation.

II. Serial Correlation and Bias in Analysts' Forecast Errors: Eight-Month Forecasts

We test for the bias and serial correlation in analysts' forecast errors by estimating:

$$FEPR_{it} = a_0 + a_1 FEPR_{it-1} + a_2 Ret_{it-1} + e_{it}, \qquad (2)$$

---

[6] In this section, we do not control for possible cross-sectional dependencies. Thus, these significance levels must be viewed with caution.

[7] The maximum value is 120.5 and the minimum value is −815.0. The first and third quartiles are −0.51 and 0.23, respectively.

Table 2
Bias and Serial Correlation in Analysts' Eight-Month Forecast Errors

($N=5{,}365$ Firm-Years; Dependent Variable$=FEPR_{it}$)

		White Method		Bootstrap Method		
Independent Variable	*OLS Coefficient*	*Standard Error*	*t-statistic*	*Coefficient*	*Standard Error*	*t-statistic*
Intercept	−1.99*	0.38	5.16*	−1.58	0.38	4.13*
$FEPR_{it-1}$	0.40*	0.17	2.30**	0.64	0.18	3.61*
Adjusted $R^2=0.13$; Chi-square Coefficient$=4.08$						
Intercept	−2.71*	0.43	6.31*	−2.53	0.42	6.10*
$FEPR_{it-1}$	0.39*	0.17	2.26**	0.36	0.17	2.10**
Ret_{it-1}	10.05*	1.27	7.94*	9.57	1.32	7.25*
Adjusted $R^2=0.15$; Chi-square Coefficient$=15.28$*						

Note: The standard errors and the t-statistics for the OLS coefficients are determined before and after adjusting for heteroscedasticity via the White (1980) method. The chi-square coefficient tests for the degree of heteroscedasticity.

$FEPR_{it}$ ($FEPR_{it-1}$)=eight-month forecast error for firm i for fiscal year t ($t-1$) divided by the stock market price at the time of the forecast; and

Ret_{it-1}=one-year stock return ending at the beginning of the month of the forecast.

* Significant at the 0.01 level (two-tailed test).

** Significant at the 0.05 level (two-tailed test).

where $FEPR_{it}$($FEPR_{it-1}$) is the analysts' eight-month forecast error for firm i for year t ($t-1$) and Ret_{it-1} is the one-year stock return ending on the month of the current forecast.

Rejecting the null that a_0 equals zero indicates that, on average, analysts' forecasts display a levels bias. Rejecting the null that a_1 equals zero shows that analysts' forecasts do not fully incorporate the information contained in past forecast errors. Rejecting the null that a_2 equals zero indicates that analysts do not fully use the past stock returns when forming expectations of future earnings. Beaver et al. (1980, 1987) and Collins et al. (1987) show that past stock returns reveal information about future earnings. If analysts' forecasts are efficient, this information should be reflected in their estimates.

Table 2 contains the pooled cross-sectional time-series regression results for equation (2), estimated with and without Ret_{it-1}. The statistical significance of the OLS coefficients are presented before and after the White (1980) adjustment for heteroscedasticity. To adjust for possible cross-sectional correlation, a bootstrapping procedure is used (Noreen 1989) to estimate the coefficients and their standard errors.[8]

The coefficient on $FEPR_{it-1}$ is significantly positive at the 0.05 level (using White's method) or 0.01 level (using OLS and bootstrap methods), indicating that forecast errors of annual earnings are correlated in adjacent years. In contrast, Givoly (1985) finds no evidence of serial correlation in annual forecast errors. One interpretation compatible with our results is that analysts systematically underestimate the permanence of past forecast errors when forecasting future earnings (see, e.g., Mendenhall 1991, 170). That is, they do not appear to properly incorporate information about the time-series properties of earnings into their forecasts.

---

[8] This procedure controls for possible cross-sectional dependence for each year's residuals. We also used yearly dummy variables to control for possible cross-sectional correlations in forecast errors, as has been done by O'Brien (1988), and Collins and Kothari (1989), who use the change in earnings as their dependent variable. The results are qualitatively the same as those reported in table 2. We do not control for time-series dependence because tests on the residuals show a lack of significant serial correlation.

The significantly negative intercept terms suggest that analysts, on average, overestimate earnings. This result supports the findings of Stober (1991) who finds that although the number of realized increases in annual *EPS* is 68 percent greater than the number of decreases, the frequency of forecasted increases outnumbers forecasted decreases by a margin of three to one. It is also consistent with the findings of Barefield and Comiskey (1975) and Fried and Givoly (1982).

O'Brien (1988) notes that pooled cross-sectional, time-series regressions may, ex *post*, falsely reject the hypothesis that analysts' forecasts are unbiased. For example, an adverse macroeconomic shock could lead to a significantly negative forecast error when, in fact, analysts may have efficiently set their forecasts.

We examine this possibility by re-estimating equation (2) on a year-by-year basis to see if the bias and serial correlations reported in table 2 are driven by a few years. We find that the intercept term is significantly negative at the 0.01 level for nine of the 12 sample years. Similarly, the coefficient on the previous year's forecast error is significantly different from zero at the 0.01 level for every year except 1989, when it is significant at the 0.05 level. The mean (median) values for the intercept and for the coefficient on the previous year's forecast error are −1.20 (−1.05) and 0.38 (0.36), each significantly different from zero.[9] Thus, the observed bias and serial correlation from the pooled regression are not specific to any one year (or a subset of years).

The bottom section of table 2 shows that the coefficient on Ret_{it-1} is significantly positive at the 0.01 level for all three statistical methods.[10] This suggests that analysts are not using past stock returns efficiently when setting their current annual forecasts, a result consistent with Abarbanell (1990), who documents a similar finding for quarterly earnings forecasts. Further, since Ret_{it-1} represents the information available to market participants, our finding is consistent with Abdel-khalik (1990), who finds that analysts' forecasts are not good proxies for market earnings expectations. It also supports Brown et al. (1987b) who find that analysts' forecasts and past stock returns together provide a better proxy of the market's expectations of future earnings than analysts' forecasts alone.

III. Analysts' Forecasts and the Time-Series Properties of Earnings

Cross-sectional Variation in Bias and Serial Correlation of Forecast Errors

In this section, we explicitly examine whether the bias and positive serial correlation shown in table 2 are observed across the full sample, or are more likely to occur when the previous period's earnings are predominantly permanent or transitory. To examine this question, we estimate:

---

[9] The t-statistic is the mean value divided by the standard deviation of the 12 yearly coefficients. A sign rank test confirms that the number of positive coefficients is different from a random selection of 12 positive/negative draws.

[10] To check if past returns are indeed related to future earnings, we regress EPS (normalized by price) for time *t* on Ret_{it-1}. The coefficient on Ret_{it-1} is 3.71, significant at the 0.01 level (with both OLS and White standard errors).

$$FEPR_{it}=b_0+b_1PERM_{it-1}+b_2FEPR_{it-1}+b_3(FEPR_{it-1}{*}PERM_{it-1})+b_4Ret_{it-1}+u_{it}, \quad (3)$$

where $PERM_{it-1}$ is a dummy variable equal to one if last year's change in earnings for firm i is deemed to be relatively permanent and zero if it is relatively transitory.[11]

To differentiate between permanent and transitory earnings components, we rely on the results presented in Beaver and Morse (1978), Ou and Penman (1989b), and in Jaffe et al. (1989), showing that earnings/price (E/P) ratios proxy for the extent to which the previous year's earnings are transitory. In particular, they show that firms with relatively high, low, or negative E/P ratios experience mean reversion in the following year's earnings, whereas firms in the middle range experience fairly stationary earnings.

Accordingly, firms are ranked each year by their current E/P ratios, which are calculated by using the previous year's earnings per share and stock price at the beginning of the month of the eight-month forecast. Ten portfolios are formed. Portfolio 1 consists of negative E/P ratios (Ou and Penman 1989b; Jaffe et al. 1989); portfolios 2–10 are formed by ranking positive E/P firms from low to high and then dividing the ranked sample into nine portfolios of equal size. Re-ranking firms each year by their current E/P ratios is equivalent to assuming that the permanence of earnings for a firm is not stationary over time (e.g., see Freeman and Tse 1989; Ou and Penman 1989b).

We define the permanent group ($PERM_{it-1}=1$) as those firm-years in portfolios 4–8 and the transitory group ($PERM_{it-1}=0$) as those in portfolios 1–3 and in 9–10.[12] We further divide the temporary group into two subgroups: firm years with negative E/P ratios (portfolio 1) and firm-years with positive, but relatively small or large E/P ratios (portfolios 2, 3, 9, and 10). Segregating negative E/P ratios is based on Jaffe et al. (1989), who show that these firms display different firm and stock return characteristics from positive E/P firms, and Ou and Penman (1989b), who present evidence that firm-years with negative E/P ratios display dramatic mean reversion.[13]

We also estimate:

$$\begin{aligned}FEPR_{it}=c_0+c_1PERM_{it-1}+c_2NEG_{it-1}+c_3FEPR_{it-1}+c_4(FEPR_{it-1}{*}PERM_{it-1})\\+c_5(FEPR_{it-1}{*}NEG_{it-1})+c_6Ret_{it-1}+u_{it}, \quad (3a)\end{aligned}$$

---

[11] Although prior studies present evidence that annual earnings, in general, follow a random walk (e.g., Albrecht et al. 1977; Ball and Watts 1972; Watts and Leftwich 1977), other studies show that, on a year-to-year basis, earnings have both permanent and transitory (i.e., mean-reverting) components (e.g., Brooks and Buckmaster 1976; Ou 1990; Ou and Penman 1989a, 1989b).

[12] To examine whether these classifications are reasonable, we examine the time-series properties of each class of firms. Specifically, we estimate: $CEPSPR_{it}=a_0+a_1CEPSPR_{it-1}+e_{it}$ for each group, where $CEPSPR_{it}$ ($CEPSPR_{it-1}$) is the change in annual EPS for firm i deflated by price for year t ($t-1$). For the permanent group, we expect a_1 to be indistinguishable from zero. For the temporary group, we expect mean-reversion in earnings, i.e., a_1 to be significantly negative.

For the permanent group, a_1 equals 0.047, insignificantly different from zero (two-tailed test) at the 0.10 level. For the temporary group, a_1 equals -0.354, significantly different from zero at the 0.01 level. The Pearson and Spearman correlations between $CEPSPR_{it}$ and $CEPSPR_{it-1}$ for the permanent group are 0.03 and 0.07; for the temporary group, they are -0.32 and -0.16. Thus, E/P appears to be a reasonable instrument for differentiating between permanent and temporary earnings.

[13] Regressing current changes in earnings on last year's change in earnings for the subsample of negative E/P firms yields a slope coefficient of -0.407, significant at the 0.01 level. This confirms that, on average, firm-years with negative earnings experience a larger degree of mean-reversion than firm-years with positive temporary earnings (see footnote 12).

Table 3
Cross-Sectional Variation in Bias and Serial Correlation in Analysts' Eight-Month Forecast Errors

(N=5,305 Firm-Years; Dependent Variable=$FEPR_{it}$)

Independent Variable	*Equation (3) (Permanent vs. Temporary Earnings Groups)*		*Equation (3a) (Permanent vs. Positive Temporary and Negative Earnings Groups)*	
Intercept	−1.94	−2.46	−1.82	−2.36
	(14.35)*	(18.32)*	(13.06)*	(17.09)*
$PERM_{it-1}$	1.21	1.32	1.09	1.23
	(6.54)*	(7.37)*	(5.78)*	(6.77)*
NEG_{it-1}			−2.88	−2.16
			(4.87)*	(3.76)*
$FEPR_{it-1}$	0.23	0.21	0.34	0.32
	(15.80)*	(14.84)*	(11.40)*	(10.88)*
$FEPR_{it-1}$*$PERM_{it-1}$	0.42	0.37	0.31	0.27
	(8.30)*	(7.61)*	(5.41)*	(4.85)*
$FEPR_{it-1}$*NEG_{it-1}			−0.24	−0.21
			(6.03)*	(5.40)*
Ret_{it-1}		5.44		5.36
		(18.73)*		(18.49)*
Adjusted R^2	0.09	0.14	0.09	0.15

Note: The numbers in parentheses are *t*-statistics.

$FEPR_{it}$ ($FEPR_{it-1}$)=eight-month forecast error for firm *i* for fiscal year *t* (*t*−1) divided by the stock market price at the time of the forecast;

$PERM_{it-1}$=dummy variable equal to 1 if the firm's previous period's earnings is predominantly permanent, and 0 otherwise;

NEG_{it-1}=dummy variable equal to 1 if the firm's previous period's earnings is negative, and 0 otherwise; and

Ret_{it-1}=one-year stock return ending at the beginning of the month of the forecast.

* Significant at the 0.01 level (two-tailed test).

where NEG_{it-1} is a dummy variable equal to one if the E/P ratio portfolio is one (the negative E/P group) and zero otherwise.

The regression results for equations (3) and (3a) are contained in table 3, which shows that the overestimation bias (intercept and dummy intercept terms) in analysts' forecasts and the serial correlation between successive years (slope and dummy slope coefficients) are not homogeneous among groups.[14]

First, the significantly negative intercept terms and the significantly positive coefficients on $PERM_{it-1}$ in both equations indicate that the level of overestimation bias is less when the previous period's earnings are predominantly permanent.[15] The significantly negative coefficient on NEG_{it-1} in equation (3a) implies that the degree of overestimation is most evident for firms with negative earnings. This latter result is consistent with Klein (1990), who finds that analysts tend to overpredict EPS most for the group of firms that recently experienced the most negative stock returns (many of which also had negative earnings).

---

[14] For these and the remaining tests reported, we eliminated outliers in which the absolute value of the forecast errors or change in earnings for year *t* or *t*−1 (scaled by price) is greater than 100 percent. This reduced the sample from 5,365 to 5,305 observations. To determine the sensitivity of our results to the presence of outliers, we used the SAS subcommand INFLUENCE to remove extreme observations systematically from the data and then re-estimated the equations with the purged data. The results with the purged data are qualitatively the same as those reported and are available from the authors.

Next, the coefficients on both $FEPR_{it-1}$ and $FEPR_{it-1}*PERM_{it-1}$, are significantly positive in equation (3), which suggests that analysts systematically omit information contained in last year's forecast error for all firms. Moreover, the tendency to omit this information is greater when previous period earnings are predominantly permanent, and less so when such earnings are predominantly transitory. The significantly negative coefficient on $FEPR_{it-1}*NEG_{it-1}$ in equation (3a), coupled with the evidence that firms with negative earnings display the most transitory earnings (see footnote 13), lends further support to this conclusion.[16] Thus, the degree of bias and serial correlation of analysts' forecast errors does not appear to be homogeneous across firm-years. Finally, the significantly positive coefficient on Ret_{it-1} implies that analysts do not completely utilize the information about future earnings contained in past stock returns.

Adaptive Expectations Model

Previous empirical studies using adaptive expectations models for analysts' forecasts support the view that analysts use this last period's forecast error to formulate the present period's forecast (see, e.g., Abdel-khalik and Espejo 1978; Brown and Rozeff 1979; Givoly 1985). However, these studies do not explicitly examine whether analysts incorporate the degree of permanence of earnings in last year's forecast errors in making their forecasts.[17]

The adaptive expectations model for analysts' forecasts of annual EPS is:

$$(FPR_{it}-FPR_{it-1})=d_0+d_1FEPR_{it-1}+v_{it}, \qquad (4)$$

where FPR_{it} (FPR_{it-1}) is the eight-month forecast for year t $(t-1)$, and $FEPR_{it-1}$ is the forecast error for year $t-1$. All variables are deflated by stock price at the beginning of the month of the eight-month forecast for year $t-1$. According to the adaptive expectations hypothesis, the more permanent the last year's earnings, the closer d_1 will be to one. Conversely, the more transient the last year's earnings, the closer d_1 will be to zero.

To examine if analysts recognize the time-series properties of earnings when setting their forecasts, we estimate:

$$(FPR_{it}-FPR_{it-1})=f_0+f_1PERM_{it-1}+f_2FEPR_{it-1}+f_3(FEPR_{it-1}*PERM_{it-1})+v_{it} \qquad (5)$$

---

[15] To determine if the permanent group exhibits significant negative bias, we test separately the null hypotheses that the sum of the intercept term and the coefficient on $PERM_{it-1}$ in the first and third columns of table 3 equals zero. We reject each null at the 0.01 level, and conclude that analysts tend to overestimate earnings for the permanent group.

[16] For example, adding the coefficients in the third column of table 3 on $FEPR_{it-1}$ and $FEPR_{it-1}*NEG_{it-1}$ produces a serial correlation coefficient of 0.10 (significant at the 0.05 level) for the negative group. In contrast, the serial correlation for the permanent group (the sum of $FEPR_{it-1}$ and $FEPR_{it-1}*PERM_{it-1}$) is 0.65. The latter value is significantly greater than the values for the positive temporary and negative groups.

[17] For example, Givoly (1985, 383) finds that the adaptation coefficient for more than 60 percent of his companies was below one and concludes that this ". . . may imply that the earnings process of most companies is perceived by analysts as being different from a random walk." However, he does not test for this.

Table 4
Adaptive Expectation Model of Setting Eight-Month Forecasts
(N=5,305 Firm-Years; Dependent Variable=$FPR_{it}-FPR_{it-1}$)

Independent Variable	*Equation (5) (Permanent vs. Temporary Earnings Groups)*	*Equation (5a) (Permanent vs. Positive Temporary and Negative Earnings Groups)*
Intercept	0.35 (4.71)*	0.32 (4.43)*
$PERM_{it-1}$	0.44 (4.34)*	0.47 (4.80)*
NEG_{it-1}		−1.30 (4.25)*
$FEPR_{it-1}$	0.27 (33.30)*	0.60 (38.49)*
$FEPR_{it-1}*PERM_{it-1}$	0.38 (13.73)*	0.04 (1.44)
$FEPR_{it-1}*NEG_{it-1}$		−0.48 (22.92)*
Adjusted R^2	0.25	0.33

Note: The numbers in parentheses are *t*-statistics.

FPR_{it} (FPR_{it-1})=eight-month forecast of year-end EPS for firm i for year t ($t-1$) divided by the stock market price at the time of the previous period's forecast;

$FEPR_{it}$ ($FEPR_{it-1}$)=eight-month forecast error for firm i for fiscal year t ($t-1$) divided by the stock market price at the time of the forecast;

$PERM_{it-1}$=dummy variable equal to 1 if the firm's previous period's earnings is predominantly permanent, and 0 otherwise; and

NEG_{it-1}=dummy variable equal to 1 if the firm's previous period's earnings is negative, and 0 otherwise.

* Significant at the 0.01 level (two-tailed test).

and

$$(FPR_{it}-FPR_{it-1})=g_0+g_1PERM_{it-1}+g_2NEG_{it-1}+g_3FEPR_{it-1}$$
$$+g_4(FEPR_{it-1}*PERM_{it-1})+g_5(FEPR_{it-1}*NEG_{it-1})+v_{it}, \tag{5a}$$

where the variables are defined as before. If analysts utilize the time-series properties of earnings, then f_3 and g_4, the slope-dummy coefficients for the permanent group, should be significantly positive. Further, g_5, the coefficient on the slope-dummy for the negative group should be significantly negative.

The regression results for equations (5) and (5a) are presented in table 4. In equation (5), the coefficient on $FEPR_{it-1}*PERM_{it-1}$ is significantly positive, indicating that d_1, the adaptation coefficient in equation (4), is significantly greater when previous period earnings are predominantly permanent. This supports the view that analysts at least partially recognize the difference between permanent and transitory earnings components. In equation (5a), the coefficient on $FEPR_{it-1}*NEG_{it-1}$ is significantly negative, signifying that analysts correctly perceive that the earnings of negative E/P firms are the least permanent. However, the coefficient on $FEPR_{it-1}*PERM_{it-1}$ in equation (5a), is not significant at the 0.05 level, suggesting that analysts may not materially differentiate the degree of permanence of earnings among firms in the permanent and the positive temporary groups. One explanation for this finding is that analysts underestimate the permanence of earnings for firms in the permanent group much more than they do for firms in the positive temporary group. This explanation is consistent with the results shown in table 3; the serial correlation of forecast errors is greater for the permanent group than for the positive temporary group.

Taken together, tables 3 and 4 suggest that analysts recognize some of the time-series properties of earnings when making forecasts. However they do not fully recognize the temporary versus permanent components of previous period earnings when forecasting future earnings.

IV. Adjusting Analysts' Forecasts of Earnings

To address the economic significance of our findings, estimates of bias and serial correlation in forecast errors are used to adjust existing forecasts, and improvements in the adjusted forecasts are examined. Specifically, we use the following mechanical procedure to adjust analysts' forecasts of earnings. First, the model given by equation (3a), without Ret_{it-1}, is estimated separately for each year. Second, simple averages of the regression coefficient estimates for five consecutive years are computed. The five-year mean coefficient is used because of the non-stationarity of coefficients found for our yearly regressions.[18] These mean estimates are then used to adjust the following year's forecast. For example, the data from 1978 to 1982 is used to estimate the coefficients. We then use the mean coefficients to adjust the actual forecasts of 1983 earnings made by analysts. This procedure is repeated with data from 1979 to 1983 to obtain the coefficients to adjust the forecasts of 1984, and so on.[19]

Table 5 reports the results for the adjustments in analysts' forecasts. Panel A presents a comparison of the forecast bias and serial correlation in forecast errors associated with the unadjusted and adjusted analysts' forecasts. Since we require five years of leading data, the number of firm-years is reduced to 3,892, spanning from 1983 through 1989. For these sample years, the bias (-1.66, $t=13.9$) and the serial correlation (0.26, $t=15.9$) are significant for the unadjusted forecasts. The adjusted analysts' forecasts exhibit very little bias (0.17, $t=1.5$) and serial correlation (0.03, $t=2.2$). Thus, the mechanical procedure yields a significant reduction in both bias and serial correlation.

In panel B, we compare the accuracy, as measured by the mean squared error, of the unadjusted and adjusted forecasts. The mean squared error for the adjusted forecasts is 12.2 percent less than for the actual forecasts. Further, the improvement in accuracy is similar across firm type; the mean squared error is reduced by 15.3, 9.9, and 12.9 percent for firms in the permanent, the positive temporary, and the negative groups, respectively.[20] We, therefore, conclude that our findings of bias and serial cor-

---

[18] We also use one-year coefficients and, as expected, find similar but weaker results.

[19] Conroy and Harris (1987) and Stickel (1990) use alternative methods to update analysts' forecasts. The first study uses a simple average of analysts' and time-series forecasts. The second updates individual analyst forecasts by using the information in consensus forecasts.

[20] An interesting result in table 5 is that the mean squared errors for the unadjusted analysts' forecasts differ across groups. They are lowest for the permanent group (26.39), followed by the positive temporary group (63.74), and highest for the negative group (238.74). A similar pattern of mean squared errors is observed for the

Table 5
Effect of a Mechanical Adjustment to Analysts' Forecasts of Earnings—Eight-Month Forecasts
(N=3,892 Firm-Years)

Panel A. Change in Bias and Serial Correlation:

Analysts' Forecast (Dependent Variable=$FEPR_{it}$)		*Adjusted Analysts' Forecast (Dependent Variable=$AFEPR_{it}$)*	
Intercept	−1.66 (13.90)*	Intercept	0.17 (1.50)
$FEPR_{it-1}$	0.26 (15.90)*	$AFEPR_{it-1}$	0.03 (2.20)**
Adjusted R^2	0.061	Adjusted R^2	0.001

Panel B. Change in Accuracy (Mean Squared Error):

	Analysts' Forecast	*Adjusted Analysts' Forecast*	*Percentage Change*
All firms	59.15	51.96	−12.2
Permanent group	26.39	22.35	−15.3
Positive temporary group	63.74	57.46	−9.9
Negative group	238.74	207.96	−12.9

Note: The numbers in parentheses are *t*-statistics.

$FEPR_{it}$ ($FEPR_{it-1}$)=eight-month forecast error for firm i for year t ($t-1$) divided by the stock market price at the time of the forecast; and

$AFEPR_{it}$ ($AFEPR_{it-1}$)=adjusted forecast error for year t ($t-1$) divided by the stock market price at the time of the forecast.

* Significant at the 0.01 level (two-tailed test).

** Significant at the 0.05 level (two-tailed test).

relation are economically significant (rather than being period-specific statistical oddities) in that they allow us to improve significantly the accuracy of analysts' eight-month forecasts significantly by using past time-series properties of earnings.

V. One-Month Forecast Errors

Empirical evidence suggests that analysts' ability to predict year-end EPS improves as the forecast horizon shrinks (see, e.g., Brown et al. 1987a; Elton et al. 1984; Klein 1990; O'Brien 1988). To examine the effect of a decrease in the forecast horizon on the informational efficiency of analysts' forecasts, we repeat all of the above analyses for one-month forecasts.

Table 6 contains the bias and serial correlation for one-month forecast errors and the coefficients for the adaptive expectations model. A total of 4,530 firm-years are available on IBES. The results provide evidence that bias, serial correlation, and

---

adjusted analysts' forecasts. One possible explanation for this result is that smaller firms are classified more often in the positive temporary and negative groups and that earnings for smaller firms can be predicted with less accuracy. We leave further investigation of this issue for future research.

Table 6
Analysis of One-Month Forecast Errors
(N=4,530 Firm-Years)

Independent Variable	*Bias and Serial Correlation in Forecast Errors (Dependent Var.=$FEPR_{it}$)*		*Adaptive Expectation Model of Setting Forecasts (Dependent Var.=$FPR_{it}-FPR_{it-1}$)*	
Intercept	−0.76	−0.64	−0.07	−0.69
	(7.32)*	(5.83)*	(0.45)	(4.49)*
$PERM_{it-1}$	0.24	0.14	0.33	1.08
	(1.81)***	(0.98)	(1.64)	(5.22)*
NEG_{it-1}		−1.03		5.89
		(3.32)*		(13.44)*
$FEPR_{it-1}$	0.09	0.11	0.18	0.29
	(5.96)*	(3.95)*	(8.35)*	(13.33)*
$FEPR_{it-1}*PERM_{it-1}$	0.12	0.10	0.33	0.29
	(2.40)**	(1.70)***	(4.31)*	(6.87)*
$FEPR_{it}*NEG_{it-1}$		−0.07		−0.45
		(1.92)***		(12.06)*
Ret_{it-1}	0.56	0.51		
	(2.30)**	(2.10)**		
Adjusted R^2	0.01	0.02	0.03	0.09

Note: The numbers in parentheses are t-statistics.

FPR_{it} (FPR_{it-1}) = one-month forecast of year-end EPS for firm i for year t ($t-1$) divided by the stock market price at the time of the previous period's forecast;

$FEPR_{it}$ ($FEPR_{it-1}$) = one-month forecast error for firm i for fiscal year t ($t-1$) divided by the stock market price at the time of the forecast;

$PERM_{it-1}$ = dummy variable equal to 1 if the firm's previous period's earnings is predominantly permanent, and 0 otherwise;

NEG_{it-1} = dummy variable equal to 1 if the firm's previous period's earnings is negative, and 0 otherwise; and

Ret_{it-1} = one-year stock return ending at the beginning of the month of the forecast.

* Significant at the 0.01 level (two-tailed test).

** Significant at the 0.05 level (two-tailed test).

*** Significant at the 0.10 level (two-tailed test).

adjusted R^2 levels are dramatically less than those found for the sample of eight-month forecast errors (table 3). Nevertheless, even at this late juncture, analysts' forecasts exhibit overprediction bias and serial correlation, as evidenced by the significantly negative intercept terms and the significantly positive coefficients on $FEPR_{it-1}$. Further, the pattern of underestimation for the one-month forecasts is similar to the sample of eight-month forecasts in that the coefficient on $FEPR_{it-1}*RM_{it-1}$ is significantly positive, whereas the coefficient on $FEPR_{it-1}*NEG_{it-1}$ is significantly negative.[21] Results similar to those in table 3 are also reported for the intercept dummy variables $PERM_{it-1}$ and NEG_{it-1}.

The regression results for the adaptive expectations model for one-month forecasts in table 6 show that analysts partially recognize the time-series properties of earnings

---

[21] The same criterion as in section IV is used to categorize firm-years into permanent, temporary positive, and negative groups.

Table 7
Effect of a Mechanical Adjustment to Analysts' Forecasts of Earnings—One-Month Forecasts
(N=3,184 Firm-Years)

Panel A. Change in Bias and Serial Correlation:

Analysts' Forecast (Dependent Variable=$FEPR_{it}$)		*Adjusted Analysts' Forecast (Dependent Variable=$AFEPR_{it}$)*	
Intercept	−0.73 (8.30)*	Intercept	0.23 (2.50)**
$FEPR_{it-1}$	0.09 (5.30)*	$AFEPR_{it-1}$	0.03 (1.60)
Adjusted R^2	0.009	Adjusted R^2	0.001

Panel B. Change in Accuracy (Mean Squared Error):

	Analysts' Forecast	*Adjusted Analysts' Forecast*	*Percent Change*
All firms	24.29	25.09	3.3
Permanent group	12.55	12.47	−0.6
Positive temporary group	22.52	23.20	3.0
Negative group	105.47	112.15	6.3

Note: The numbers in parentheses are *t*-statistics.

$FEPR_{it}$ ($FEPR_{it-1}$)=one-month forecast error for firm *i* for year *t* (*t*−1) divided by the stock market price at the time of the forecast; and

$AFEPR_{it}$ ($AFEPR_{it-1}$)=adjusted forecast error for year *t* (*t*−1) divided by the stock market price at the time of the forecast.

* Significant at the 0.01 level (two-tailed test).

** Significant at the 0.05 level (two-tailed test).

when setting short-term forecasts. This is demonstrated by the significantly positive coefficients on $FEPR_{it-1}*PERM_{it-1}$ and the significantly negative coefficient on $FEPR_{it-1}*NEG_{it-1}$. Thus, similar conclusions about analysts recognizing the degree of permanence for these respective groups can be made for both one-month and eight-month forecasting horizons.

Table 7 examines whether the observed serial correlations and biases documented in table 6 can be used to improve the accuracy of short-term forecasts. The same procedure employed for adjusting the longer-term forecasts is used. Panel A shows that the adjusted forecasts display less bias and serial correlation in forecast errors than the unadjusted forecasts.

Panel B reports the change in the mean squared error for the one-month forecasts. As the horizon shrinks from eight months to one, analysts show considerable improvement in their forecasting ability, evident in the reduction of the mean squared error for unadjusted forecasts from 59.15 (table 5) to 24.29. However, the adjustment procedure does not improve the accuracy of analysts' one-month forecasts any further. For the full sample, the mean squared error actually increases by 3.3 percent, and no significant increase in accuracy is found for any of the individual groups. We attribute this latter result to the low explanatory power (R-squared values) of the regressions of one-month forecast errors on prior forecast errors. Thus, even though the adjustment procedure leads to reductions in bias and serial correlation, it does not improve the accuracy of the short-term forecasts.

VI. Summary and Conclusions

In this study, we examine whether analysts correctly use the time-series properties of annual earnings when setting their forecasts of annual EPS. We document significantly positive serial correlation in eight-month and one-month forecast errors, a result consistent with the hypothesis that analysts, on average, underestimate the permanence of the last year's forecast error when setting forecasts. We also find a significant overprediction bias for both forecast horizons.

Further examination of our results reveals that the observed bias and serial correlation of the prediction errors are not uniform across firm-years. Specifically, the over-prediction bias is most pronounced for firms that previously reported negative annual earnings. Also, the positive serial correlation in the prediction errors is greatest for firms with previous period earnings that are predominantly permanent.

To evaluate the economic significance of our findings, we use our estimates of bias and serial correlation to adjust existing forecasts. The adjusted eight-month forecasts are found to be more accurate predictors of future earnings than the actual forecasts made by analysts. In addition, the adjustment procedure leads to a significant reduction in the bias and serial correlation for both the one-month and eight-month forecast errors.

References

Abarbanell, J. 1990. Analysts' earnings forecasts as market expectations around stock price changes. Working paper, University of Michigan, Ann Arbor.

———, and V. L. Bernard. 1991. Do analysts underreact or overreact to earnings? Working paper, Harvard University, Cambridge, MA.

Abdel-khalik, A. R. 1990. Specification problems with information content of earnings: Revisions and rationality of expectations and self-selection bias. *Contemporary Accounting Research* 7 (Fall): 142–72.

———, and J. Espejo. 1978. Expectations data and the predictive value of interim reporting. *Journal of Accounting Research* 16 (Spring): 1–13.

Albrecht, W. S., L. L. Lookabill, and J. C. McKeown. 1977. The time series properties of annual earnings. *Journal of Accounting Research* 15 (Autumn): 226–44.

Ball, R., and R. Watts. 1972. Some time-series properties of accounting income. *Journal of Finance* 27 (June): 663–82.

Barefield, R. M., and E. E. Comiskey. 1975. The accuracy of analysts' forecasts of earnings per share. *Journal of Business Research* 3 (July): 241–51.

Beaver, W. H., and D. Morse. 1978. What determines price-earnings ratios? *Financial Analysts Journal* 34 (July/August): 65–76.

———, R. A. Lambert, and D. Morse. 1980. The information content of security prices. *Journal of Accounting & Economics* 2 (March): 3–28.

———, ———, and S. G. Ryan. 1987. The information content of security prices: A second look. *Journal of Accounting & Economics* 9 (July): 139–57.

Bernard, V. L., and J. Thomas. 1990. Evidence that stock prices do not fully reflect the implications of current earnings for future earnings. *Journal of Accounting & Economics* 13 (December): 305–40.

Brooks, L., and D. Buckmaster. 1976. Further evidence of the time series properties of accounting income. *Journal of Finance* 31 (December): 1359–73.

Brown, L. D., R. L. Hagerman, P. A. Griffin, and M. E. Zmijewski. 1987a. Security analyst superiority relative to univariate time-series models in forecasting quarterly earnings. *Journal of Accounting & Economics* 9 (April): 61–87.

———, ———, ———, and ———. 1987b. An evaluation of alternative proxies for the market's assessment of unexpected earnings. *Journal of Accounting & Economics* 9 (July): 159–93.

———, and M. Rozeff. 1979. Adaptive expectations, time-series models and analyst forecast revision. *Journal of Accounting Research* 17 (Autumn): 341–51.

Collins, D. W., and S. P. Kothari. 1989. An analysis of intertemporal and cross-sectional determinants of earnings response coefficients. *Journal of Accounting & Economics* 11 (July): 143–81.

———, ———, and J. D. Rayburn. 1987. Firm size and the information content of prices with respect to earnings. *Journal of Accounting & Economics* 9 (July): 111–38.

Conroy, R., and R. Harris. 1987. Consensus forecasts of corporate earnings: Analysts' forecasts and time-series methods. *Management Science* 33 (June): 725–38.

Elton, E., M. Gruber, and M. Gultekin. 1984. Professional expectations: Accuracy and diagnosis of errors. *Journal of Financial and Quantitative Analysis* 19 (December): 351–63.

Freeman, R., and S. Tse. 1989. The multi-period information content of earnings announcements: Rational delayed reactions to earnings. *Journal of Accounting Research* 27 (Supplement): 49–79.

Fried D., and D. Givoly. 1982. Financial analysts' forecasts of earnings—A better surrogate for market expectations. *Journal of Accounting & Economics* 4 (October): 85–107.

Givoly, D. 1985. The formation of earnings expectations. *The Accounting Review* 60 (July): 372–86.

Jaffe, J. J., D. B. Keim, and R. Westerfield. 1989. Earnings yields, market values and stock returns. *Journal of Finance* 44 (March): 135–48.

Klein, A. 1990. A direct test of the cognitive bias theory of share price reversals. *Journal of Accounting & Economics* 13 (July): 155–66.

Mendenhall, R. 1991. Evidence of possible underweighting earnings-related information. *Journal of Accounting Research* 29 (Spring): 170–79.

Noreen, E. 1989. *Computer-Intensive Methods for Testing Hypotheses: An Introduction*. New York: John Wiley & Sons.

O'Brien, P. 1988. Analysts' forecasts as earnings expectations. *Journal of Accounting & Economics* 10 (January): 53–83.

Ou, J. A. 1990. The information content of non-earnings accounting numbers as earnings predictors. *Journal of Accounting Research* 28 (Spring): 144–63.

———, and S. H. Penman. 1989a. Financial statement analysis and the prediction of stock returns. *Journal of Accounting & Economics* 11 (November): 295–329.

———, and ———. 1989b. Accounting measurement, price-earnings ratio, and the information content of security prices. *Journal of Accounting Research* 27 (Supplement): 111–44.

Penman, S. H. 1987. The distribution of earnings news over time and seasonalities in aggregate stock returns. *Journal of Financial Economics* 18 (June): 199–228.

Smith, D. B., and S. Pourciau. 1988. A comparison of the financial characteristics of December and non-December year-end companies. *Journal of Accounting & Economics* 10 (December): 335–44.

Stickel, S. E. 1990. Predicting individual analyst earnings forecasts. *Journal of Accounting Research* 28 (Autumn): 409–17.

Stober, T. L. 1991. Summary financial statement measures and analysts' forecasts of earnings. Working paper, Indiana University, Bloomington, IN.

Watts, R. L., and R. W. Leftwich. 1977. The time series properties of annual accounting earnings. *Journal of Accounting Research* 15 (Autumn): 253–71.

White, W. H. 1980. A heteroskedasticity-consistent covariance matrix estimation and a direct test for heteroskedasticity. *Econometrica* 48 (May): 817–38.

THE ACCOUNTING REVIEW
Vol. 67, No. 1
January 1992
pp. 199–211

Competition and Audit Fees

Michael W. Maher
University of California–Davis
Peter Tiessen
University of Alberta
Robert Colson
Case Western Reserve University
Amy J. Broman
Michigan State University

SYNOPSIS AND INTRODUCTION: This study describes the behavior of audit fees during a period of apparent increasing competition in the market for independent audit services. Academic researchers (e.g., Danos and Eichenseher 1986; Kinney 1988) and the business press (e.g., *Journal of Accountancy* 1984; *The Wall Street Journal* 1985, 1987; Work 1985) have noted increasing competition in the market for audit services among public accounting firms, but there has been no documentation that audit fees have decreased. The purpose of this study is to determine whether real audit fees decreased between 1977 and 1981. This interval started with federal investigations of anticompetitive behavior in the accounting profession, included several changes in the accounting profession that, coupled with the economic downturn of the late 1970s and early 1980s, could have increased competition in the market for audit services. It ended after the Federal Trade Commission announced it was closing its investigation of the profession's anticompetitive rules because many of the restrictions against competition had been dropped (*The Wall Street Journal* 1980).

The study sample was developed from a set of publicly traded companies reporting external audit fees in a University of Michigan database.

We are grateful to the Paton Accounting Center, University of Michigan for providing access to the data. We wish to thank Betsy MacLean and Charles Klemstine for their research assistance. We are grateful to Vic Bernard, Nick Dopuch, Jere Francis, William Kinney, Katherine Schipper, Dan Simunic, Miriam Wells, the two referees, and participants in workshops at the University of North Carolina, Case-Western Reserve University and The Ohio State University for their helpful comments.

Submitted January 1987 and December 1990.
Accepted June 1991.

Of the 98 companies reporting these data for 1977 and 1981, the two years of interest, 20 were excluded to minimize the potential confounding effects of major changes in internal auditing, the impact of auditor turnover, and changes in regulation of the banking industry on external audit fees.

The study finds a significant decrease in real audit fees between 1977 and 1981. These findings were not sensitive to alternative specifications of the audit fee model and were not driven by any particular industry or audit firm. The results of this study are consistent with claims of increasing fee competition in the market for independent audit services. In view of the number of changes occurring in the audit profession and the market during that period, it is difficult to make causal inferences about the effects of particular changes in the profession on audit fees. Thus, this article should be viewed as a descriptive study of the behavior of audit fees in a time when the market for audit services was allegedly becoming more competitive.

Key Words: *Auditing, Audit fees, Economics of auditing, Accounting profession.*

Data Availability: *Data used in this study will be made available subject to restrictions to maintain the confidentiality of the participating companies.*

THE next section of this paper presents the research methods and describes the sample. The following section describes the empirical results. The final section discusses conclusions and limitations.

I. Research Methods

The tests examine the change in audit fees between 1977 and 1981. In 1977, the baseline year, the Federal Trade Commission (The *Wall Street Journal* 1977) and the United States Justice Department (Kohn 1977; Lowe 1987) began investigations of anticompetitive behavior in the market for audit services, and the Metcalf Committee of the United States Senate published its report criticizing the accounting profession for anticompetitive behavior (U.S. Senate 1977). During the 1977–81 period, the American Institute of Certified Public Accountants repealed Rule 401 of the Code of Professional Ethics (now called the Code of Professional Conduct), effective March 31, 1979, which had prohibited client encroachment, and it modified rules prohibiting client solicitation and advertising (AICPA 1979, 1981; Lowe 1987). In 1980, the Federal Trade Commission closed its probe of the accounting profession because many of the anticompetitive restrictions had been dropped. The test year, 1981, was selected to allow sufficient time for a reaction in the audit services market to an apparent increase in competition. We acknowledge, however, that this is a relatively short time interval for the effects to be observed.

The analyses use matched-pairs first-differences designs in which each auditee company is its own control before and after the events of interest. Although we control

Table 1
Sample Selection Procedure

Category of Companies	Number
Initial Sample of Companies Reporting External Audit Fee Data for 1977 and 1981	98
Banks and Other Financial Institutions	14
Auditees with an Auditor Change Between 1974 and 1981	4
New Internal Audit Department or Substantial Reorganization of Internal Audit Department Between 1974 and 1981	2
Final Sample	78

explicitly for particular audit fee covariates known to affect the level of audit fees, there remains a potential for omitted variable bias.

Sample

We developed the sample using publicly available data from SEC 10-K reports, *Standard and Poor's Register of Corporations, Directors and Executives*, and *Who Audits America*, as well as data on external audit fees and internal audit costs from a database at the Paton Accounting Center, University of Michigan. The Michigan database,[1] developed from a nonrandom sample of companies who agreed to participate in a study of internal auditing, contained 98 publicly traded companies reporting external audit fees for 1977 and 1981. This sample was reduced by the following restrictions (summarized in table 1) to mitigate the confounding effects of changes in the competitive environment of auditing and other factors affecting audit fees.

First, we excluded banks and other financial institutions for two reasons. This was a period of major changes in the banking industry, including deregulation, that potentially affected the audit fee structure of banks. Further, previous research (Simunic 1980) has shown a significant industry effect for banking, implying that the audit fee function is different in banks than in other companies. This restriction reduced the sample by 14 companies.

Second, the sample is restricted to companies that did not change auditors in the previous three years to avoid confounding the effects of changes in the competitive environment with such auditor change effects as "low-balling" pricing strategies (DeAngelo 1981; Francis and Simon 1987) or type of contracts (Palmrose 1989). This restriction reduced the sample by another four companies.

Third, in view of Wallace's (1984,199) estimate that "most internal audit departments allocate about 5 percent of their time to assisting external auditors," we further restricted the sample to reduce the effect of internal audit changes on external audit fees. Participants in a pilot study comprised of external auditors, internal auditors, and financial executives were asked to describe the internal/external audit relation and to identify conditions under which increases in internal auditing would reduce external audit fees. They stated that substitution of internal audit for external audit is likely to occur when new internal audit departments are installed or when there is a substantial improvement in an internal audit department that has previously not been sufficiently competent for reliance by external auditors. After an internal audit department reaches a threshold size and level of competence, however, additional internal audit resources would have virtually no effect on external audit fees, according to the pilot study participants. Consequently, the sample is restricted to auditees who had established an internal audit department before 1974, which is three years before the baseline year, and who had made no major changes in the organization of their internal audit department, such as centralizing a previously decentralized department, between 1974 and 1981. This reduced the sample by two more companies.

---

[1] See Mautz et al. (1984) for detailed information about how the Michigan database was developed.

The resulting sample is a set of mature, relatively stable companies, each audited by the same audit firm during the period. Although not specifically intended in the research design, these companies were all audited by Big Eight firms. As shown in table 2, these are relatively large companies covering a range of industries outside of financial services.

Inferences from this study should note that this is not a random sample. The database from which our sample is drawn was comprised of companies voluntarily participating in a study by the University of Michigan. Further, we selected only the companies reporting external audit fees in 1977 and 1981, then reduced that sample as described.

Data Reliability

We contacted the companies that had reported external audit fees for 1977 and 1981 in the University of Michigan database to verify the accuracy of the audit fee data and to obtain information about the companies' internal audit functions. The reported audit fees were reviewed with the respondents to assure that audit fees were not confounded with other CPA firm billings, such as consulting or tax billings. We were able to isolate audit fees from other CPA firm billings for all 78 companies studied. We also obtained information about the companies' internal audit functions to eliminate companies that had started or substantially changed their internal audit departments, as discussed above.

II. Findings

The first set of data compares the annual audit fee for the ith auditee in 1981 with its 1977 audit fee. The 1981 audit fee is restated in 1977 dollars with the Department of Labor Bureau of Labor Statistics Consumer Price Index.[2] On the basis of interviews with audit partners, we concluded that payments of audit fees are generally made over a several-month period falling on either side of the auditee's fiscal year-end. Virtually all of the companies had December 31 fiscal year-ends, so we used the December 1981 Consumer Price Index to restate 1981 fees in 1977 dollars. We refer to audit fees restated in this manner as "real" audit fees henceforth.

---

[2] We checked the sensitivity of the results to the use of a particular price index and found the GNP deflator, the Producer Price Index, and particular industry indexes gave virtually identical results to those using the Consumer Price Index.

Table 2
Descriptive Statistics

Variables	*1977* *Mean* *(Standard Deviation)*	*1981* *Mean* *(Standard Deviation)*
External Audit Fee[a]	$762,760 (900,290)	$1,045,400 (1,248,800)
Assets[a]	$2,248 million (3,641 million)	$3,439 million (5,579 million)
Revenues[a]	$2,534 million (5,167 million)	$3,814 million (6,560 million)
Two-digit SIC Codes	3.35 (3.45)	3.30 (2.75)
Subsidiaries[5]	4.15 (2.60)	4.12 (2.48)
Ratio of Foreign to Total Assets	.14 (.16)	.14 (.16)
Ratio of Receivables to Total Assets	.16 (.11)	.18 (.11)
Ratio of Inventories to Total Assets	.18 (.16)	.17 (.15)

Companies	*Number of Companies*	*Percent*
Industry Breakdown		
Manufacturing:		
Food Products (20)	5	6.4
Chemicals (28)	8	10.3
Petroleum Refining (29)	4	5.1
Rubber and Miscellaneous Plastics (30)	3	3.8
Stone, Clay, Glass, and Concrete (32)	3	3.8
Electronic and Electrical Equipment (36)	7	9.0
Transportation Equipment (37)	7	9.0
All Other Manufacturing[b]	14	18.0
Total Manufacturing	51	65.4
Mining (12, 13, 14)	3	3.8
Utilities (49)	12	15.4
Wholesale and Retail Trade (51, 53)	4	5.1
Insurance (63, 64)	5	6.4
Services (73, 78, 89)	3	3.8
Total Companies	78	100.0
Audited by Big Eight Firms	78	100.0
No Auditor Turnover, 1974–81	78	100.0

Note: The numbers in parentheses are two-digit SIC codes.

[a] Dollar amounts are nominal dollars.

[b] Includes companies drawn from nine other two-digit SIC industries. No other manufacturing two-digit SIC code industry had more than two companies.

Table 3
Descriptive Statistics on the Change in External Audit Fees between 1977 and 1981

	Means (Standard Deviations)			
	1977	*1981*	*Amount of Change*	*Percent Change*
External Audit Fee	$762,760 (900,290)	$691,880 (826,500)	$70,880 decrease	9.3 decrease
Revenues	$2,534 million (5,167 million)	$2,567 million (4,402 million)	$33 million increase	1.3 increase
Assets	$2,248 million (3,641 million)	$2,308 million (3,742 million)	$60 million increase	2.7 increase

Note: 1981 amounts are adjusted to 1977 dollars by using the Bureau of Labor Statistics Consumer Price Index. With 1977 as a base=1.000, the 1981 indexes were 1.513 for December-to-December price level changes (used for external audit fees) and 1.490 for average 1977 to average 1981 price changes (used for assets and revenues).

Table 3 presents the means and standard deviations of audit fees and auditee size (in 1977 dollars) for 1977 and 1981. The data show that mean real audit fees decreased 9.3 percent during the period. Since size has been found to be a significant covariate with audit fees in prior research (see e.g., Francis 1984; Palmrose 1986) and explains nearly 50 percent of the cross-sectional variation in levels of audit fees in our sample, we checked to see whether the decrease in real audit fees can be attributed to a decrease in auditee size adjusted for price level changes. To adjust size for price level changes, we use the average Consumer Price Index for 1981 to restate measures of size into 1977 dollars. We found the average real size of auditees in the sample actually increased slightly, using either revenues (1.3 percent increase) or assets (2.7 percent increase) as the measure of size, thus ruling out size decrease as the cause of the audit fee decrease.

Univariate Tests

We first ascertained whether real audit fees decreased between 1977 and 1981 for a significant number of auditees by using univariate tests. Table 4 shows that real audit fees decreased for 51 of the 78 auditees studied, which is significantly more than half of the sample, using a binomial test.

We also did the binomial test after controlling for size, which, as noted above, is an important covariate in cross-sectional estimates of audit fees. Although both assets and revenues have been used to measure size in previous audit fee studies, we use revenue to measure the change in auditee size. This avoids potential measurement problems that we found in some companies in which step increases in assets occur when old, largely depreciated assets are replaced by new assets. Our sensitivity analysis indicated no difference in results when we replaced assets with revenues to measure size.

After controlling for changes in size, table 4 shows that real audit fees decreased for 56 of the 78 auditees in the sample, which is significantly more than half, with a binomial test. To test the sensitivity of the results to our measure of auditee size and to

Table 4
Change in Audit Fees: Univariate Tests

Fee Changes	*Binomial Test*	*Wilcoxon Signed-Ranks Test*
Change in External Audit Fees (1981 fee/Deflator[a] − 1977 fee)		
Auditees with Decreases	51*	
Auditees with Increases	27	
Rank Sum of Decreases		2,126**
Rank Sum of Increases		955
Change in Ratio of External Audit Fees to Size $\left(\frac{\text{1981 fee/Deflator}}{\text{1981 revenue/Deflator}} - \frac{\text{1977 fee}}{\text{1977 revenue}}\right)$:		
Auditees with Decreases	56*	
Auditees with Increases	22	
Rank Sum of Decreases		2,416**
Rank Sum of Increases		665

[a] The deflator is from the Bureau of Labor Statistics Consumer Price Index.
* Significantly greater than one-half at $p<0.01$ with a binomial test (one-tailed).
** Significant decrease at $p<0.01$ with the Wilcoxon matched-pairs signed-ranks test (one-tailed).

the assumed linear relation between change in audit fee and change in auditee size implied in these tests, we also performed the tests using assets as the measure of auditee size, and the natural logarithms of audit fee and size, and found the results to be unaffected by these alternative measures.

Table 4 also shows the results of the Wilcoxon signed-ranks test. The rank sum of audit fee decreases is significant for differences in audit fees alone ($p<0.01$) and for differences in audit fees adjusted for differences in size ($p<0.01$). Again, substituting assets for revenues as the measure of size and using natural logarithms of audit fees and measures of size did not affect the results.

Multivariate Analysis

Prior cross-sectional research has consistently shown that the complexity of auditees and the proportion of auditee assets that are "risky" (e.g., inventories and receivables) are significant covariates with audit fees (Colson et al. 1988; Elliot and Korpi 1978; Francis 1984; Francis and Simon 1987; Palmrose 1986; Simon 1985; Simunic 1980). As noted earlier, size is also a significant covariate in cross-sectional estimates of audit fees. This study differs from previous studies that examined the levels of audit fees because we examine first-differences in audit fees. It is reasonable to assume, nevertheless, that changes in size, complexity, and risky balance sheet assets would affect changes in audit fees, given the rationale for their effect on levels of fees.[3] Consequently, we next control for the effects of these covariates.[4]

---

[3] The rationale for these effects is as follows. The greater the auditee size, the greater will be the number of elements that comprise the accounting population and thus the greater the sample size required to achieve a given level of control. The greater the complexity through decentralization and diversification of the financial reporting entity, the greater will be the number of decision centers in an organization whose activities need to be monitored. The greater the proportion of "risky" balance sheet assets, the greater the auditor's loss exposure and need for special audit steps (e.g., confirmation of receivables, observation of inventories).

Using revenue as the measure of size for the reason previously discussed, we examine the effect of auditee complexity and the risky balance sheet assets on audit fees, using measures similar to prior research (see, e.g., Simunic 1980). Using a matched-pairs first-differences design, as described in the model below, this test indicates whether audit fees decreased after controlling for auditee size, complexity, and risky balance sheet assets. The estimated model is:

$$FEE_i = b_o + b_1 REV_i + b_2 DIVERS_i + b_3 SUBS_i^{.5} + b_4 FORGN_i + b_5 INV_i + b_6 RECV_i + \hat{e}_i, \tag{1}$$

where:

FEE_i = external audit fees in 1981 minus external audit fees in 1977 for the *i*th auditee; 1981 fees are converted to 1977 dollars by using the end-of-year Consumer Price Index;

REV_i = revenue in 1981 minus revenue in 1977 for the *i*th auditee; 1981 revenue is converted to 1977 dollars by using the Consumer Price Index average for the year;

$DIVERS_i$ = change in the number of two-digit SIC codes between 1977 and 1981 for the *i*th auditee;

$SUBS_i^{.5}$ = square root of the number of consolidated subsidiaries in 1981 minus the square root of the number of consolidated subsidiaries in 1977 for the *i*th auditee;

$FORGN_i$ = change in the ratio of foreign to total assets between 1977 and 1981 for the *i*th auditee;

INV_i = change in the ratio of inventories to total assets between 1977 and 1981 for the *i*th auditee;

$RECV_i$ = change in the ratio of receivables to total assets between 1977 and 1981 for the *i*th auditee.

If audit fees decreased after controlling for these covariates, then we expect to find $b_0 < 0$. Table 5 presents the results of estimating the model in equation (1).[5] These estimates control for changes in auditee size, risky balance sheet assets, and complexity that could affect the change in audit fees. The intercept, b_0, is negative, as predicted, and significant ($p < 0.01$).

Discussion

The findings are consistent with claims of increasing competition in the market for audit services. There are several changes in the marketplace and the accounting profession that could explain these results. For example, changes in the Code of Professional Ethics, particularly the changes allowing for greater client solicitation, may have reduced the transactions costs of switching auditors along the lines suggested by DeAngelo (1981). Increasing client participation in auditing (Palmrose 1989) and

---

[4] We replicated Simunic's (1980) model to see whether his cross-sectional estimate of audit fee levels holds for our 1977 and 1981 samples. The results for our sample were almost identical to Simunic's large auditee sample and are presented in the appendix.

[5] Our diagnostic tests, including examining residual plots, tests for heteroscedasticity using the Goldfeld-Quandt test and the Kolmogorov test for normally distributed residuals, did not reveal significant differences from the assumptions of the regression model.

Table 5
Change in Audit Fees: Multivariate Tests

Variables		*Coefficient* *(Standard Error)*
Intercept	b_0*	−66,358** (23,268)
REV	b_1	0.0000702*** (0.0000143)
DIVERS	b_2	32,734*** (13,294)
SUBS[a]	b_3	−31,419 (20,454)
FORGN	b_4	−592,340 (441,270)
INV	b_5	32,069 (465,240)
RECV	b_6	−397,980 (418,170)
$n=78$		
Adjusted R^2	0.31	
F-statistic	6.62**	

* Predicted to be negative.
** Significant at $p \leq 0.01$ (one-tailed test).
*** Significant at $p \leq 0.01$ (two-tailed test).

increased documentation of internal controls following the Foreign Corrupt Practices Act (Maher 1981) could also help companies reduce real audit fees. Coupled with the downturn in the economy, these factors could enable companies, particularly large, publicly traded companies like those in our sample, to bargain effectively with auditors to reduce real audit fees.

Sensitivity Analysis

We re-estimated equation (1) by replacing revenues with assets as the measure of size to check the sensitivity of our results to the measure of size, which is important because auditee size is such a significant determinant of audit fees. The results were virtually the same with respect to both the overall fit of the regression and the sign and significance of b_o. We checked the sensitivity of the results to the first-differences model by running a levels model with 1981 as an indicator variable. The coefficient on the indicator variable was negative and significant at $p \leq 0.05$ ($t=-1.87$).

We also checked for sensitivity of the results to linear versus nonlinear changes in the fee/size relationship. Prior cross-sectional studies of the determinants of audit fee levels have found an improved fit by using nonlinear fee/size relations, such as the log of fees regressed on the log of auditee size. This *ad hoc* fit for determinants of audit fee *levels* would not necessarily apply to our estimate of first-differences in audit fees; nevertheless, we re-estimated equation (1) using the first-differences of the log of fees and the log of size measures to check the sensitivity of our findings to nonlinear estimates of the fee/size relation. The basic findings regarding the sign and significance of b_0 were not sensitive to the use of logs, but the regression fit was not as good (adjusted $R^2=0.14$ for the model with log-revenues and adjusted $R^2=0.13$ for the model with log-assets). These results imply that rejection of the null hypothesis is not sensitive to the use of logs; however, the regression fit for the estimate of changes in audit fees is not as good when using log transformations.

Other Factors Affecting Audit Fees

Changes in Internal Auditing. We checked whether the decrease in external audit fees could be explained by the substitution of increased internal audit resources for external services. From interviews with internal and external auditors, we learned that substitution of internal for external service is likely to occur when new internal audit departments are installed, but after an internal audit department reaches a threshold size and level of competence, additional internal audit resources have virtually no effect on external audit fees. By selecting a sample of large auditees that had well-established internal audit departments throughout the period of study, we expected to eliminate the effects of changes in internal resources on changes in external audit fees. One company that had started a new internal audit department and one that had substantially reorganized its internal audit department after 1974 were dropped from the initial sample of 98 companies, as shown in table 1. The questionnaires and follow-up telephone interviews with the auditees in this study did not reveal any other companies with substantial changes in internal audit departments that the auditees believed would affect the amount of external audit work done for the period 1977–81. In addition, we tested for the possible substitution of internal audit costs for external audit fees by adding a variable to equation (1) to control for internal audit cost changes from 1977–81. The coefficient of the internal audit variable was not significant ($t=-0.02$). The t-statistic on the coefficient of the regression intercept in equation (1) decreased from 2.85 without controlling for internal audit to 2.73 with the internal audit variable in equation (1).[6]

Audit Fee Trends. Changes in the competitive environment of auditing began as early as 1972 (Kinney 1988), which could result in a long-term downward trend in audit fees. To check for this trend, we compared the rate of audit fee decrease during the period of our study, 1977–81, with a previous control period, 1974–77. These rates of growth or decline were standardized for changes in size, by using revenues, and converted to 1974 dollars by using the Consumer Price Index. The sample size is smaller

---

[6] The simultaneous nature of the relation between internal and external audit costs calls for two-stage least squares regression (Johnston 1984, 442–77). First, we obtain predicted values of internal audit costs from two cross-sectional regressions (one for each year, 1977 and 1981) of internal audit costs on the following independent variables. The reduced-form internal audit model was:

$$ICOST_i=b_0+b_1REV_i+b_2DIVERS_i+b_3SUBS_i+b_4FORGN_i+b_5INV_i+b_6RECV_i+b_7BRIBE_i+\hat{e}_{it}$$

where *ICOST* is the amount of internal audit salaries paid by the auditee, and *BRIBE* is an indicator variable with a value of one if the company had reported paying bribes during the SEC investigations preceding the Foreign Corrupt Practices Act. The other variables have been previously defined, but they are measured in levels in this model instead of first-differences. Second, we include the change in predicted values of internal audit cost as an independent variable in the model estimating the change in audit fees.

for this comparison ($n = 63$) than for the 1977–81 tests because less data were available for 1974. Also, we did not control for risky balance sheet assets and complexity variables because of data limitations.

Annualizing the rates of decline for exposition purposes, we found audit fees decreased 1.365 percent (in real terms) per year during the period 1974–77, and 3.908 percent per year during the test period 1977–81 (for the same set of 63 companies). These data indicate a downward trend in audit fees throughout the entire period, 1974–81; however, the decrease in fees was marginally statistically significantly greater for the 1977–81 period than for the 1974–77 period ($t = 1.374$, $p \leq 0.10$, one-tailed test). These results imply a long-term decrease in real audit fees, possibly because of the 1972 change in the Code of Professional Ethics (Kinney 1988).

Industry Effects. As noted earlier, we excluded banking from the study because the changes in the banking industry could confound the effects of changes in competition in the market for audit services. We knew of no potential confounding for any other particular industry, *ex ante*, that would cause it to be excluded from the sample. Consequently, we simply checked to see if our results were driven by any of the industries in our sample, using indicator variables for two-digit SIC codes of industry membership. None of the coefficients on these indicator variables was significant.

Audit Firm Effects. We also checked to see if the results were driven by any particular audit firm by using indicator variables for the audit firms. None of the coefficients was significant, indicating that no particular audit firm was driving the results. We also classified audit firms according to their audit technology into "structured," "unstructured," and "intermediate" groups (Kinney 1986) and found that, although audit fees decreased for all three groups, the fee decreases were greater for unstructured firms than for either intermediate or structured firms.[7] While these potential differences between groups of audit firms is noted, no interpretation is made as to the effects of audit firm structure on changes in audit fees.

III. Conclusions and Limitations

The purpose of this article was to describe audit fee behavior during a period in which the market for audit services presumably became more competitive. We found a significant decrease in audit fees between 1977 and 1981, which is consistent with claims that the audit industry has become increasingly competitive.

We selected a sample designed to reduce confounding effects and used a matched-pairs first-differences design. Furthermore, we controlled for the three variables known in the literature to be covariates with audit fees, namely, auditee size, auditee complexity, and risky assets. We found the results to be insensitive to various measures of size and functional form.

Nevertheless, the study has several limitations. First, the results should be viewed as descriptive in view of our inability to control for various changes in the audit profession and the marketplace during the period of this study. Second, the literature contains little knowledge about the "true" functional form of changes in audit fees. In particular, the effects of omitted variables like changes in the documentation of internal controls that could affect the scope of audit work, the impact of changes in client participation in audits (aside from internal auditing), changes in the auditor's risk assessment and attitude toward risk, change in the economic conditions of particular audit firm offices, and changes in the aggressiveness of particular audit firms or offices, among other factors, could affect changes in audit fees. Third, the audit fee data are self-reported and, although verified by follow-up telephone calls, are subject to error. Fourth, the sample of auditees that provided both audit fee data and publicly available data was limited and not randomly selected. These and other limitations should be considered in making inferences from this study.

---

[7] Of the Big Eight firms, unstructured firms were Coopers & Lybrand and Price Waterhouse; structured firms were Deloitte, Haskins & Sells, Touche Ross (now combined into Deloitte and Touche), and Peat Marwick Mitchell (now part of KPMG Peat Marwick Main). Arthur Young, Ernst & Whinney (now combined into Ernst & Young) and Arthur Andersen & Co. were intermediate structure firms.

Appendix

Impact of Complexity on Fee/Assets[5]: Cross-Sectional Tests Comparing Our Sample with Simunic's (1980) Large Auditee Sample

		Our Sample		*Simunic's Sample*
Variables	*Predicted Sign*	*1977 Coefficient (Standard Error)*	*1981 Coefficient (Standard Error)*	*1977 Coefficient (Standard Error)*
Intercept	+ or −	1.68 (1.99)	.41 (2.19)	5.14 (6.30)
Two-digit SIC Codes	+	.93* (.30)	.79* (.39)	.72* (.31)
Subsidiaries[5]	+	.22 (.46)	.80* (.48)	.93* (.18)
Ratio of Foreign to Total Assets	+	23.42* (7.49)	25.29* (7.25)	14.61* (2.88)
Receivables to Total Assets	+	36.26* (8.67)	36.38* (9.13)	18.93* (4.46)
Inventories to Total Assets	+	23.57* (5.83)	27.56* (6.90)	9.09* (2.98)
n		78	78	202
Adjusted R^2		.61	.62	.51
F		25.3	26.5	18.4

Note: Simunic (1980) also included variables measuring auditee profitability; the number of years the auditee used the auditor; whether the auditee was a utility, received a "subject to" opinion, or was audited by Price Waterhouse versus one of the other Big Eight firms. None of the coefficients on these variables was statistically significant for his large auditee sample. Simunic's large auditee sample included only companies with sales greater than $125 million, which is more comparable to our sample than his total sample.

* Significant at $p \leq 0.05$ (one-tailed test).

References

American Institute of Certified Public Accountants. 1979. *Referendum: Background Information on Proposed Amendments to the Code of Ethics.* New York: AICPA.

———. 1981. *Bylaws, Rules of Conduct.* New York: AICPA.

Colson, R., M. Maher, A. Broman, and P. Tiessen. 1988. Audit pricing models for regulation research: Reliability and stability. *Research in Accounting Regulation* 2 (March): 61–79.

Danos, P., and J. Eichenseher. 1986. Long-term trends toward seller concentration in the U.S. audit market. *The Accounting Review* 61 (October): 633–50.

DeAngelo, L. 1981. Auditor independence, "low balling" and disclosure regulation. *Journal of Accounting & Economics* 3 (August): 113–27.

Elliot, R., and A. Korpi. 1978. Factors affecting audit fees. In *Cost-Benefit Analysis of Auditing, Commission on Auditors' Responsibility Research Study No. 3,* edited by M. Shakun. New York: AICPA.

Francis, J. 1984. The effect of audit firm size on audit prices: A study of the Australian market. *Journal of Accounting & Economics* 6 (August): 133–51.

———, and D. T. Simon. 1987. A test of audit pricing in the small-client segment of the U.S. audit market. *The Accounting Review* 62 (January): 145–57.

Johnston, J. 1984. *Econometric Methods.* 3d ed. New York: McGraw-Hill.

Journal of Accountancy. 1984. Pressures on accounting profession share focus of address by Commissioner Treadway. 158 (December): 14–16.

Kinney, Jr., W. 1986. Audit technology and preferences for auditing standards. *Journal of Accounting & Economics* 8 (March): 73–89.

———. 1988. Attestation research opportunities: 1987. *Contemporary Accounting Research* 4 (Spring): 416–25.

Kohn, I. 1977. The professions in a changing world. *The CPA Journal* 47 (November): 7–9.

Lowe, H. J. 1987. Ethics in our 100-year history. *Journal of Accountancy* 163 (May): 78–87.

Maher, M. 1981. The impact of regulation on controls: Firms' response to the foreign corrupt practices act. *The Accounting Review* 56 (October): 751–70.

Mautz, R., P. Tiessen, and R. Colson. 1984. *Internal Auditing: Directions and Opportunities.* Altamonte Springs, FL: Institute of Internal Auditors.

Palmrose, Z. 1986. Audit fees and auditor size: Further evidence. *Journal of Accounting Research* 24 (Spring): 97–110.

———. 1989. The relation of audit contract type to audit fees and hours. *The Accounting Review* 64 (July): 488–99.

Simon, D. 1985. The audit services market: Additional empirical evidence. *Auditing: A Journal of Practice & Theory* 5 (Fall): 71–78.

Simunic, D. 1980. The pricing of audit services: Theory and evidence. *Journal of Accounting Research* 18 (Spring): 161–90.

United States Senate. 1977. *The Accounting Establishment.* Washington, D.C.: 95th Cong., 1st sess. (March 31).

The Wall Street Journal. 1977. FTC to study state accountancy boards, groups for possible curbs on competition. (March 25): 7.

———. 1980. FTC closes probe of accounting rules limiting competition. (September 17): 4.

———. 1985. Audit fees fall as CPA firms jockey for bids. (January 28): 33.

———. 1987. Marketing battles erupt among business-hungry CPA firms. (May 21): 1.

Wallace, W. 1984. *A Time Series Analysis of the Effect of Internal Audit Activities on External Audit Fees.* Altamonte Springs, FL: Institute of Internal Auditors.

Work, C. 1985. Accounting's bottom line: Big trouble. *U.S. News and World Report* (October 21): 58.

THE ACCOUNTING REVIEW
Vol. 67, No. 1
January 1992
pp. 212–215

A Comment on the Proper Use of Residuals in Patell's Squared Unexpected Return Statistic

William M. Cready
Texas A & M University

IN this comment I employ the term residual to mean a difference between an observation's actual value and its expected value when the observation's actual value is also employed in deriving its expected value. A residual is distinguished from a forecast error in that a forecast error equals the difference between an observation's actual value and its expected value when the observation's actual value is *not* employed in deriving its expected value. This distinction between residuals and forecast errors is important because tests of squared unexpected returns as proposed by Patell (1976) are based on forecast errors, not on residuals.

In the Patell tests, residuals are measured during "estimation" or "non-event" periods that do not include the information event under consideration, to provide estimates of the parameters of the return distributions, such as the α and β terms of the market model and especially the variance of the disturbance term. These estimates then are used to assess the significance of forecast errors (the differences between actual and expected returns) on the information event dates. Patell (1976) derives the mean and variance of several test statistics under the assumption of normality, and Marais (1984) extends this analysis to include situations in which the kurtosis of the stock return distribution exceeds that of the normal distribution.

A recent article by Stice (1991) illustrates the danger of failing to distinguish between residuals and forecast errors in calculating normal-theory squared unexpected return statistics in the manner proposed by Patell. Stice does not separate the return data into estimation and event periods. Rather, he uses the event period observations, measured as residuals, both to estimate the variance of the return distribution and to measure "abnormal" price activity on the event dates. Stice then finds that, over the joint estimation/event period, a Patell-type three-day squared unexpected return averages 0.850, rather than its normal-theory value of one. In the analysis that follows, I demonstrate that this 0.850 average occurs entirely because Stice employs residuals in his analysis.

I would like to thank two anonymous referees for their comments.

Submitted February 1991.
Accepted July 1991.

Stice calculates the squared unexpected return (SR) for firm i over the three-day period t as:

$$SR_{it}=\left[\left[\frac{T-3}{T-1}\right]^{1/2}(r_{it}-\bar{r}_i)\bigg/\left[\sum_{t=1}^{20}(r_{it}-\bar{r}_i)^2/(T-1)\right]^{1/2}\right]^2; \tag{1}$$

where:

r_{it} is the three-day return on firm i's stock;
$\bar{r}_i$ is the average three-day return on firm i's stock over the 20 estimation/event period three-day return observations; and
T is the number of estimation/event period observations (20 in this case).

This expression simplifies to:

$$SR_{it}=(T-3)(r_{it}-\bar{r}_i)^2\bigg/\sum_{t=1}^{20}(r_{it}-\bar{r}_i)^2. \tag{2}$$

The expected value of SR_{it} can then be computed by summing SR_{it} across t and averaging to give:

$$E(SR_{it})=(1/T)(T-3). \tag{3}$$

In Stice T equals 20 so $E(SR_{it})=0.85$ for all i, which is the exact value Stice reports as his estimation-period sample wide mean squared standardized unexpected return.

Table 1 reports Stice's results adjusted for the correct squared residual mean of 0.85.[1] Relative to the Z-values reported by Stice, the adjusted Z-values do not take on large negative values in non-event period windows. Furthermore, the significance levels of these adjusted Z's correspond quite closely to the significance levels based on Stice's empirical distribution analysis.[2]

In concluding, I think two additional points as to how this comment pertains to the Stice analysis are noteworthy. First, this comment in no way undermines Stice's conclusion that the earnings price response occurs at the earnings announcement date rather than at the earlier 10-K or 10-Q date. Indeed, it actually strengthens his findings; after adjusting Stice's squared return results for the correct mean (0.85), the finding of an average squared return of 1.101 at the earnings announcement date becomes highly significant under the normal-theory test as well as the empirical-distribution test. Second, Stice's analysis suggests that there is at least one advantage to employing residuals when examining squared returns rather than forecast errors. Specifically, research by Patell (1976), Marais (1985), and Cready and Mynatt (1990, 1991) find that means of squared return prediction errors differ from their normal-theory-derived

---

[1] I am indebted to one of the anonymous referees for providing this table.

[2] The mean-adjusted Z-values reported in table 1 are still incorrect since they are based on the analytically derived variance of the squared standardized prediction errors (i.e., $2(T-2)/(T-5)$), which equals the variance of the squared standardized residuals only asymptotically. The use of the prediction error variances is necessary because I do not possess the data needed to calculate a sample-based estimator of this variance, and the lack of independence among the residuals makes it impossible to analytically derive the variance of the squared standardized residuals in a manner similar to Patell's derivation of the distribution and variance of squared standardized prediction errors.

Table 1
Comparison of Reported Z-Values with Adjusted Z-Values With Values Derived from Stice (1991, table 2)

			Significance Level	
Three-Day Interval	*Reported Z-Value*	*Adjusted Z-Value*	*Adjusted Z-Value*	*Reported Randomization*
SEC_{-27}–SEC_{-25}	−0.686	1.105	0.135	0.139
SEC_{-24}–SEC_{-22}	−2.619	−0.828	0.796	0.802
SEC_{-21}–SEC_{-19}	−0.922	0.869	0.193	0.195
SEC_{-18}–SEC_{-16}	−2.107	−0.316	0.624	0.625
SEC_{-15}–SEC_{-13}	−0.631	1.160	0.123	0.130
SEC_{-12}–SEC_{-10}	−2.382	−0.591	0.723	0.725
SEC_{-9}–SEC_{-7}	−3.231	−1.440	0.925	0.948
SEC_{-6}–SEC_{-4}	−2.757	−0.967	0.833	0.850
SEC_{-3}–SEC_{-1}	−2.074	−0.284	0.612	0.609
SEC_{0}–SEC_{+2}	−2.923	−1.133	0.871	0.890
WSJ_{-1}–WSJ_{+1}	1.208	2.999	0.001	0.002
WSJ_{+2}–WSJ_{+4}	−0.178	1.612	0.053	0.058
WSJ_{+5}–WSJ_{+7}	−1.966	−0.176	0.570	0.563
WSJ_{+8}–WSJ_{+10}	−2.178	−0.388	0.651	0.646
WSJ_{+11}–WSJ_{+13}	−1.660	0.130	0.448	0.442
WSJ_{+14}–WSJ_{+16}	−1.860	−0.070	0.528	0.524
WSJ_{+17}–WSJ_{+19}	−2.394	−0.604	0.727	0.729
WSJ_{+20}–WSJ_{+22}	−2.195	−0.405	0.657	0.651
WSJ_{+23}–WSJ_{+25}	−1.475	0.316	0.376	0.372
WSJ_{+26}–WSJ_{+28}	−2.780	−0.989	0.839	0.851

The reported Z-value calculations assumed a value of 1 for the mean of the squared residuals.

The adjusted Z-value calculations use a value of 0.85 for the mean of the squared residuals.

The adjusted Z-value significance levels are one-tailed levels based on the standard normal distribution, assuming that the variance of the squared standardized residuals is $2(T-2)/(T-5)$.

The reported randomization significance levels are one-tailed levels based on an empirical distribution generated as described in Stice (1991).

mean of 1, meaning their true expected value is unknown.[3] The Stice analysis demonstrates that no such problem of an unknown expected value arises when residuals are employed since the expected value of the squared residuals can be computed without appeal to any form-of-distribution assumption.

---

[3] Patell (1976), Marais (1984, 1985), and (most strongly) Lobo and Mahmoud (1989) report *SR* means of less than one. In a further analysis of a 1,000-firm sample, however, Marais (1985) finds that *SR* means exceed one, a result consistent with the *SR* findings reported in Cready and Mynatt (1990, 1991). The Cready and Mynatt (1990) findings are particularly relevant because they are from a replication of Lobo and Mahmoud (1989). Employing the same sample and time period as Lobo and Mahmoud, they find that *SR* means exceed one, rather than fall below one.

References

Cready, W., and P. Mynatt. 1990. The empirical distribution of squared unexpected returns: An econometric audit. Working paper, The University of North Carolina, Chapel Hill.

———, and ———. 1991. The information content of annual reports: A price and trading response analysis. *The Accounting Review* 66 (April): 291–312.

Lobo, G., and A. Mahmoud. 1989. Relationship between differential amounts of prior information and security return variability. *Journal of Accounting Research* 27 (Spring): 116–34.

Marais, L. 1984. An application of the bootstrap method to the analysis of squared, standardized market model prediction errors. *Journal of Accounting Research* 22 (Supplement): 34–53.

———. 1985. Some applications of computer intensive statistical methods to empirical research in accounting. Ph. D. dissertation, Stanford University, Palo Alto, CA.

Patell, J. 1976. Corporate forecasts of earnings per share and stock price behavior: Empirical tests. *Journal of Accounting Research* 14 (Autumn): 246–76.

Stice, E. 1991. The market reaction to 10-K and 10-Q filings and to subsequent *The Wall Street Journal* earnings announcements. *The Accounting Review* 66 (January): 42–55.

THE ACCOUNTING REVIEW
Vol. 67, No. 1
January 1992
pp. 216–218

A Reply to a Comment on the Proper Use of Residuals in Patell's Squared Unexpected Return Statistic

Earl K. Stice
Rice University

CREADY (1992) correctly points out the failure of Stice (1991) to distinguish between residuals and forecast errors in computing Z-statistics derived from squared unexpected returns. This failure had the effect of introducing a significant negative bias in the resulting Z-statistics reported by Stice. Cready shows that correcting for this error causes the significance levels for these Z-statistics to correspond closely to the empirical distribution significance levels used as the basis for the Stice results. This reply offers further information about the relative distributional properties of Z-statistics derived from squared residuals and squared forecast errors.

Distribution of Z-Statistics Derived from Squared Residuals and Squared Forecast Errors

Cready suggests that there may be an advantage to using squared residuals since their mean is fixed by construction, whereas the mean of squared forecast errors has been shown to differ from its theoretical value. The Stice data is used to provide more information on this possibility, but only 232 of the 342 firm-quarters included in the original data set are used in order to allow for acquisition of estimation period returns from CRSP. Estimation period returns are used to compute forecast errors, and the Z-statistics derived from these forecast errors are compared with those derived from residuals. For the residuals, the mean and variance of returns is computed by using the 20 three-day intervals in the test period, as in Stice. For the forecast errors, the mean and variance of returns is computed by using an estimation period of the 20 three-day intervals immediately preceding the test period. Squared residual and forecast error data are used to generate empirical distributions of Z-statistics as described in Stice.

These empirical distributions, based on 999 simulated observations, are shown in table 1. The first column shows the distribution of Z-statistics based on squared residuals computed as in Stice (equation 3) with the exception that the correct mean of 0.85 (i.e., $(T-3)/T$, where $T=20$, the number of three-day intervals in the test period) is used. The distribution is fairly symmetrical around zero but has thinner tails than does a standard normal distribution, and a Kolmogorov test rejects the hypothesis that the distribution is standard normal.

Received September 1991

Table 1

Empirical Distributions of Z-Statistics Derived from Squared Unexpected Return Measures Sample is 232 of the 342 Firm-Quarters Used in Stice (1991) Based on 999 Simulated Observations

	(1) *Residuals, Theoretical Variance*	(2) *Residuals, Empirical Variance*	(3) *Forecast Errors*	(4) *Truncated Forecast Errors*	(5) *Standard Normal*
Mean	0.016	0.013	2.849	−0.165	0.000
Standard Deviation	0.859	0.977	2.719	1.070	1.000
1st Percentile	−1.822	−2.127	−1.843	−2.462	−2.326
5th Percentile	−1.342	−1.548	−1.137	−1.978	−1.645
25th Percentile	−0.590	−0.664	0.830	−0.903	−0.674
50th Percentile	−0.006	−0.017	2.520	−0.149	0.000
75th Percentile	0.560	0.626	4.534	0.552	0.674
95th Percentile	1.491	1.687	7.707	1.613	1.645
99th Percentile	2.165	2.374	10.392	2.465	2.326

Column 1: Empirical distribution of Z-statistics derived from squared residuals. Variance of squared residuals is assumed to be $(2(T-2)/(T-5))$, where $T=20$.

Column 2: Empirical distribution of Z-statistics derived from squared residuals. Variance of squared residuals is estimated from the data.

Column 3: Empirical distribution of Z-statistics derived from squared forecast errors.

Column 4: Empirical distribution of Z-statistics derived from squared forecast errors truncated at 98th percentile value of 9.057.

Column 5: Values from a standard normal distribution corresponding to the indicated percentiles.

As mentioned by Cready (footnote 2 of previous article), the analytically derived variance used to compute the Z-statistics $[2(T-2)/(T-5)]$ is only asymptotically equal to the variance of the squared residuals. Instead of using this variance, another possibility is to estimate the appropriate variance with the observed values of the squared residuals in the test period.[1] Accordingly, firm-specific estimates of the variance in squared residuals were computed and used in place of the theoretical variance. The distribution of Z-statistics computed with this variance estimate is given in the second column of table 1. The correspondence between this distribution and the standard normal distribution in the fifth column of table 1 is quite close, and, in fact, the distribution cannot be rejected as being standard normal at typical confidence levels.

The third column of table 1 contains the characteristics of the distribution of Z-statistics based on squared forecast errors. The mean of the distribution is 2.849, which significantly exceeds its predicted value of zero. In addition, the distribution exhibits positive skewness, which suggests the existence of large outliers among the squared forecast errors. To see how the properties of the distribution are affected by outliers, the underlying squared forecast errors are truncated at the 98th percentile (9.057), and these truncated values are used to recompute the Z-statistics. This distribution is included in the fourth column of table 1. Although the distribution can be rejected as being standard normal, it is symmetrical around its mean of −0.165 and represents a significant improvement in specification over the distribution generated without truncation of outliers.

---

[1] This procedure was suggested by Professor Cready.

Note that truncation causes the mean of the Z-statistics derived from squared forecast errors to drop from above zero to below zero. This is because the mean of the underlying squared forecast errors decreases from 1.284 to 0.981. Cready (footnote 3 of previous article) indicates that past research has resulted in conflicting findings about whether squared forecast errors typically are above or below 1. Differential treatment of outliers may be a partial explanation for these conflicting results.

Summary

When the necessary mean adjustment identified by Cready is made, the distribution of Z-statistics derived from squared residuals is shown to be symmetrical around zero, in contrast to the significant negative bias originally reported by Stice. It is also shown that when an empirical estimate of the variance of squared residuals is used in place of the analytically derived asymptotic variance, the distribution of Z-statistics that results cannot be rejected as being standard normal. The Z-statistics derived from squared forecast errors are shown to exhibit a positive bias and to be skewed by large positive values. This is due to the presence of large outliers among the squared forecast errors.

Taken together, these results suggest that parametric tests based on squared residuals are better specified than those based on squared forecast errors, particularly when an empirical estimate of the variance in squared residuals is used.[2] It is interesting to note that the results reported by Stice are robust across all metrics mentioned above when significance is evaluated with the approximate randomization technique described by Noreen (1989).

[2] A disadvantage of using squared residuals is that any test period price response affects computation of the residuals, reducing the power of the tests. This limitation was pointed out by Professor Cready.

References

Cready, W. 1992. A comment on the proper use of residuals in Patell's squared unexpected return statistic. *The Accounting Review* 67 (January): 212–215.

Noreen, E. 1989. *Computer-intensive Methods for Testing Hypotheses: An Introduction.* New York: John Wiley & Sons.

Stice, E. 1991. The market reaction to 10-K and 10-Q filings and to subsequent *The Wall Street Journal* earnings announcements. *The Accounting Review* 66 (January): 42–55.

THE ACCOUNTING REVIEW
Vol. 67, No. 1
January 1992
pp. 219–229

American Accounting Association

Committees 1991–92

EXECUTIVE COMMITTEE

President—Arthur R. Wyatt, Arthur Andersen & Co.
President-Elect:—Gary L. Sundem, Accounting Education Change Commission
Vice President—Andre Zund, University of St. Gall
Vice President—David A. Wilson, Ernst & Young
Vice President—Mark A. Wolfson, Stanford University
Secretary-Treasurer—Jane O. Burns, Texas Tech University
Director of Education—Robert Ingram, University of Alabama
Director of Research—Nicholas Dopuch, Washington University
Director of Publications—Kenneth J. Euske, Naval Postgraduate School
Past President—Alvin A. Arens, Michigan State University

COUNCIL

Section and Group Representatives

Accounting, Behavior & Organizations Section—Mark E. Haskins, University of Virginia, and Howard Rockness, University of North Carolina
Administrators of Accounting Programs Group—James E. Smith, College of William & Mary
Auditing Section—Theodore J. Mock, University of Southern California, and Carl S. Warren, University of Georgia
Gender Issues in Accounting—Cheryl R. Lehman, Hofstra University
Government and Nonprofit Section—Jesse W. Hughes, Old Dominion University
International Accounting Section—Juan Rivera, University of Notre Dame, and Kathleen Sinning, Western Michigan University
IS/MAS Section—Casper E. Wiggins, Texas A&M University
Management Accounting Section—Donald L. Madden, University of Kentucky, and Robert Capettini, San Diego State University
Public Interest Section—Jere R. Francis, University of Iowa
Taxation (ATA) Section—Edward J. Schnee, University of Alabama, and John O. Everett, Virginia Commonwealth University
Two-Year College Section—Billie Cunningham, Collin County Community College

Regional Group Representatives

Mid-Atlantic Regional Group—Charles J. Pineno, Clarion University of Pennsylvania
Midwest Regional Group—Paul Fischer, University of Wisconsin and Thomas Hubbard, University of Nebraska
Northeast Regional Group—Anthony T. Krzystofik, University of Massachusetts, and Philip Jagolinzer, University of Southern Maine
Ohio Regional Group—Dennis Kimmel, University of Akron
Southeast Regional Group—Thomas C. Turner, University of North Carolina and Robert W. Rouse, Security Exchange Commission
Southwest Regional Group—Jackson A. White, University of Arkansas
Western Regional Group—James F. Sepe, Santa Clara University, and J. Hal Reneau, Arizona State University

Members-at-Large

Victor Bernard, University of Michigan
Brenda Birkett, Southern University
Julie Collins, University of North Carolina
W. Morley Lemon, University of Waterloo

REGIONAL VICE PRESIDENTS

Mid-Atlantic

Charles J. Pineno, Clarion University of Pennsylvania

Midwest

Thomas Hubbard, University of Nebraska

Northeast

Anthony T. Krzystofik, University of Massachusetts

Ohio

Dennis Kimmell, University of Akron

Southeast

Thomas C. Turner, University of North Carolina

Southwest

Jackson A. White, University of Arkansas

Western

James F. Sepe, Santa Clara University

SECTION OFFICERS

Accounting, Behavior & Organizations Section

Chairman: Mark E. Haskins, University of Virginia
Chair-Elect: Paul H. Munter, The University of Miami
Secretary-Treasurer: Don W. Finn, Texas Tech University
Immediate Past Chair: Howard Rockness, University of North Carolina

Administrators of Accounting Programs Group

President: James E. Smith, College of William & Mary
Vice President: Michael A. Diamond, University of Southern California
Secretary: Jerry E. Trapnell, Clemson University
Treasurer: Penelope J. Yunker, Western Illinois University

Taxation (ATA) Section

President: Edeward J. Schnee, University of Alabama
Vice President: Cherie J. O'Neil, University of South Florida
President-Elect: John O. Everett, Virginia Commonwealth University
Secretary-Treasurer: Debra W. Hopkins, University of Texas

Artificial Intelligence/Expert Systems

Chairperson: Ed Blocher, University of North Carolina
Chairperson-Elect: Dan O'Leary, University of Southern California
Vice-Chairperson: Carol Brown, Oregon State University
Secretary-Treasurer: Rien Bouwman, University of North Carolina

Auditing Section

Chairperson: Theodore J. Mock, University of Southern California
Vice Chairperson (Practice): Brent C. Inman, Coopers & Lybrand
Vice Chairperson (Academic): William Messier, University of Florida
Secretary-Treasurer: Jane F. Mutchler, Penn State University

Gender Issues in Accounting

Chairperson: Cheryl R. Lehman, Hofstra University
Vice Chairperson: Theresa Hammond, Boston College
Chairperson-Elect: Rita Hull, Virginia Commonwealth University
Secretary-Treasurer: Robert A. Nehmer, Georgia State University

Government and Nonprofit Section

Chairperson: Jesse W. Hughes, Old Dominion University
Chairperson-Elect: Earl Wilson, University of Missouri
Secretary: W. T. Wrege, Ball State University

Information Systems/Management Advisory Services Section

Chairperson: Casper E. Wiggins, Jr., Texas A&M University
Vice Chairperson (Practice): Rod C. Thomas, Coopers & Lybrand
Vice Chairperson (Academic): Anita S. Hollander, Florida State University
Secretary-Treasurer: Paul J. Steinbart, Memphis State University

International Accounting Section

Chairperson: Juan Rivera, University of Notre Dame
Vice Chairman (Academic): Kathleen Sinning, Western Michigan University
Vice Chairman (Practice): Keith Scott, CGA Research Foundation
Secretary: James Schweikert, University of Richmond
Treasurer: Maureen Berry, Commerce West

Management Accounting Section

President: Donald L. Madden, University of Kentucky
Vice President: L. Gayle Rayburn, Memphis State University
Secretary-Treasurer: Susan F. Haka, Michigan State University

Public Interest Accounting Section

Chair: Jere R. Francis, University of Iowa
Vice Chair: Mary S. Doucet, University of Georgia
Secretary-Treasurer: Paula B. Thomas, Middle Tennessee State University
Chair-Elect: Marilyn Neimark, CUNY-Baruch College

Teaching and Curricula

Chair: Frederick L. Neumann, University of Illinois
Vice-Chair (Academic): Jan R. Williams, University of Tennessee
Vice-Chair (Practice): Jean Wyer, Coopers & Lybrand
Secretary: Kent St. Pierre, James Madison University
Treasurer: Mary Beth Armstrong, California Polytechnic State University

Two-Year College Section

Chairman: Billie Cunningham, Collin County Community College
Vice Chairman: Patrick T. Reihing, Nassau Community College
Secretary: Linda Lessing, SUNY
Regional Program Chair: Helen Gerrard Miami University

ACCOUNTING ACCREDITATION COMMITTEE

Charges: 1. To monitor the AACSB review of accreditation and other developments in accreditation that would potentially impact accounting programs.

2. To inform the AAA Executive Committee of relevant accreditation issues and recommend policies or positions as appropriate.

Chairperson: Robert G. May, University of Texas at Austin
Kenneth Banet, Grant Thornton
David R. Campbell, Case Western Reserve University
John L. Kramer, University of Florida
Joseph A. McHugh, University of Massachusetts at Boston
Albert R. Mitchell, James Madison University
Penelope J. Yunker, Western Illinois University

ACCOUNTING EDUCATION ADVISORY COMMITTEE

Charges: 1. To develop and issue position statements on matters relating to accounting education.

2. To develop and recommend to the Executive Committee educational policy for the Association.

3. To initiate, coordinate, and administer the activities of the Association in the field of accounting education, including accounting education research and continuing professional education. Effective liaison with regions and sections is an integral part of this responsibility.

4. To coordinate the work of committees for the Doctoral Consortium, Doctoral Programs, New Faculty Consortium, Senior Faculty Consortium, Trueblood Seminars, Professional Examinations, and other education-designated committees that may be established during the year.

Chairperson: Robert W. Ingram, University of Alabama
Bruce A. Baldwin, Arizona State University-West Campus
Elba F. Baskin, Arthur Andersen & Co.
Gary L. Fish, Illinois State University
Hugh A. Hoyt, Miami University
Paul A. Janell, Northeastern University
Russell J. Peterson, University of Akron
David N. Ricchiute, University of Notre Dame
G. Fred Streuling, Brigham Young University
Milton F. Usry, University of West Florida
Thomas H. Williams, University of Wisconsin-Madison

ACCOUNTING EDUCATION CHANGE IMPLEMENTATION COMMITTEE

Charges: 1. To monitor developments of the Accounting Education Change Commission.

2. To consider alternative initiatives for the AAA that are responsive to AECC developments.

3. To recommend to the Executive Committee specific initiatives for AAA implementation.

Chairperson: Donald E. Kieso, Northern Illinois University
Lewis F. Davidson, Florida International University
Brent Inman, Coopers & Lybrand
Gary A. Luoma, University of South Carolina
Corine T. Norgaard, University of Connecticut
Joseph J. Schultz, Arizona State University
Helen Gerrard, Miami University

ASSOCIATION STAFFING EVALUATION COMMITTEE

Charges: 1. To identify Association activities currently performed by volunteers that could be more effectively performed by paid permanent staff.

2. To identify Association activities currently performed by paid permanent staff that could be more effectively contracted or performed by volunteers.

3. To recommend to the Executive Committee changes, if any, in the responsibilities of paid permanent staff.

4. To plan a strategy for the eventual succession of the Association Executive Director.

5. To recommend to the Executive Committee a financing plan for all recommendations.

Chairperson: Roger H. Hermanson, Georgia State University
Jack C. Gray, University of Minnesota
Thomas McKee, East Tennessee State University
Deborah H. Turner, Georgia Institute of Technology
Norman R. Walker, Price Waterhouse

ASSOCIATION STRUCTURE EVALUATION COMMITTEE

Charges: 1. To evaluate the existing membership, responsibilities, and selection processes of the Executive Committee, the Nominating Committee, and Council.

2. To recommend revisions to the ByLaws and/or Policies and Procedures Manual to enable the American Accounting Association to serve better the interests of its existing and potential members. Cost to the Association should be a significant consideration in restructuring recommendations.

Chairperson: Harold Q. Langenderfer, University of North Carolina
Kathleen R. Bindon, University of Alabama
Robert Ellyson, Coopers & Lybrand
Michael Gibbins, University of Alberta
Robert W. Liptak, Faison & Associates
Spencer J. Martin, University of Rhode Island
Gary John Previts, Case Western Reserve University
Wanda A. Wallace, College of William and Mary
Mark A. Wolfson, Stanford University

BYLAWS COMMITTEE

Charge: To review the Association's Bylaws and recommend appropriate revisions to the Executive Committee.

Chairperson: Spencer J. Martin, University of Rhode Island
Kenneth H. Heller, George Mason University
Harold Q. Langenderfer, University of North Carolina
William Markell, University of Delaware

CARTER SCHOLARSHIPS COMMITTEE

Charges: 1. To collect and evaluate applications for the Carter Foundation Student Scholarships.

2. To recommend amounts of stipends.

3. To rank applications and to transmit the ranking to the trustees of the Foundation.

Chairperson: N. Allen Ford, University of Kansas
Geraldine F. Dominiak, Texas Christian University
Wayne E. Leininger, Virginia Polytechnic Institute and State University

COMPETITIVE MANUSCRIPT AWARD COMMITTEE

Charge: To select one, two, or, at most, three recipient(s) of the Competitive Manuscript Award using criteria approved by the Executive Committee.

Chairperson: Ira Solomon, University of Illinois
Rick Antle, Yale University
Victor L. Bernard, University of Michigan
Stephen T. Limberg, University of Texas at Austin
Krishna G. Palepu, Harvard University
Michael D. Shields, San Diego State University

CONFERENCE, CONSORTIUM, AND SEMINAR CRITERIA COMMITTEE

Charges: 1. To recommend criteria the Executive Committee should use in deciding the nature of conferences, consortia, and seminars the Association will offer.

2. To evaluate existing Association conferences, consortia, and seminars to determine whether their purposes and content are consistent with the criteria.

3. To evaluate existing Association conferences, consortia, and seminars to determine whether policies and procedures for invitations, fees, funding, responsibilities of committees, and other matters are equitable and consistently applied.

4. To recommend changes in policies and procedures for their administration.

Chairperson: Larry E. Rittenberg, University of Wisconsin-Madison
William J. Ihlanfeldt, Shell Oil Company
William Markell, University of Delaware
R. D. Nair, University of Wisconsin-Madison
Russell M. Barefield, University of Georgia
Dennis R. Reigle, Arthur Andersen & Co.
Joanne Rockness, North Carolina State University
Gerald Smith, University of Northern Iowa
Lawrence A. Tomassini, University of Illinois

CONTINUING PROFESSIONAL EDUCATION EVALUATION COMMITTEE

Charges: 1. To consider the recommendations of the 1990–91 Continuing Professional Education Evaluation Committee and the criteria developed for evaluating professional education programs sponsored by the Association.

2. To modify or reaffirm those criteria.
3. To evaluate existing continuing education programs using the developed criteria.
4. To make additional recommendations for AAA continuing professional education program.
5. To coordinate all activities and recommendations with the Accounting Education Advisory Committee.

Chairperson: Gary L. Fish, Illinois State University
Elba F. Baskin, Arthur Andersen & Co.
Karen L. Hooks, University of South Florida
Hugh A. Hoyt, Miami University
Beverly Terry, Central Piedmont Community College

CORPORATE ACCOUNTING POLICY SEMINARS COMMITTEE

Charges: 1. To plan the 1992 Corporate Accounting Policy Seminar in accordance with established AAA policies.

2. To arrange to have the seminar conducted and administered.
3. To report on the results and effectiveness of the seminar and to recommend changes by November 1, 1992.

Chairperson: William J. Ihlanfeldt, Shell Oil Company
Eugene Flegm, General Motors Corporation
J. Edward Ketz, Penn State University
Frank C. Minter, Samford University
Robert W. Rouse, Auburn University
Michael G. Tearney, University of Kentucky
David A. Ziebart, University of Illinois
James Don Edwards, University of Georgia

DELOITTE & TOUCHE WILDMAN MEDAL AWARD COMMITTEE

Charge: To administer the John R. Wildman Medal award program in accordance with the terms of the grant from the Deloitte & Touche Foundation.

Chairperson: Eugene E. Comiskey, Georgia Institute of Technology
Nabil Hassan, Wright State University
C. Jevons Lee, Tulane University
Thomas F. Schaefer, Florida State University
Philip H. Siegel, University of Houston-Downtown
Lester M. Sussman, Deloitte & Touche
Robert D. Taylor, Florida Atlantic University

DOCTORAL CONSORTIUM COMMITTEE

Charges: 1. To plan the 1992 Doctoral Consortium in accordance with established AAA policies.

2. To arrange to have the consortium conducted and administered.
3. To report on the results and effectiveness of the consortium and to recommend changes by August 1, 1992.

Chairperson W. Steve Albrecht, Brigham Young University
Russell M. Barefield, University of Georgia
Betty R. Jackson, University of Colorado at Boulder
James Loebbecke, University of Utah
Zoe-Vonna Palmrose, University of Southern California
Terrence J. Shevlin, University of Washington

DOCTORAL FELLOWSHIP COMMITTEE

Charge: To select the 1991–1992 recipients of American Accounting Association fellowship grants using the criteria approved by the Executive Committee and to determine the amounts of these grants.

Chairperson: Stanley Biggs, University of Connecticut
Constantine Konstans, Memphis State University
Kurt J. Pany, Arizona State University

DOCTORAL PROGRAMS CONFERENCE COMMITTEE

Charges: 1. To finalize the plans for the 1992 Doctoral Programs Conference in accordance with established AAA policies.

2. To arrange to have the conference conducted and administered.

3. To report on the results and effectiveness of the conference and to recommend changes by June 1, 1992.

Chairperson: J. Hal Reneau, Arizona State University
Kenneth Austin, Memphis State University
Susan Haka, Michigan State University
Douglas A. Snowball, University of Florida
Bill N. Schwartz, Virginia Commonwealth University

EXTERNAL RESOURCES DEVELOPMENT COMMITTEE

Charges: 1. To accumulate information related to sources of external funding for existing AAA programs.

2. To consider whether it is desirable and feasible to establish annual commitment levels for external resource providers to eliminate frequency of requests.

3. To consider whether additional external resource providers can be identified.

4. To consider whether the AAA should compensate external resource providers in some way, e.g., memberships, publications, etc., for their continuing support.

Chairperson: Donald H. Skadden, American Institute of Certified Public Accountants
Charles B. Eldridge, Ernst & Young
Robert E. Holtfreter, Fort Hays State University
Robert M. McDowell, Coopers & Lybrand
Bernard J. Milano, KPMG Peat Marwick
Thomas J. Nessinger, Arthur Andersen & Co.
James G. Bartlett, Richard D. Irwin, Inc.
Larry Scott, Price Waterhouse
D. Gerald Searfoss, Deloitte & Touche
John K. Simmons, University of Florida
Jan R. Williams, University of Tennessee
Eugene Flegm, General Motors Corp.

FINANCE COMMITTEE

Charges: 1. To prepare financial goals and strategies for the Association.

2. To evaluate additional or alternative ways to finance existing Association activities.

3. To work with the 1991–1992 President-Elect in preparing the 1992–1993 budget.

4. To recommend how AAA funds should be invested in accordance with established Executive Committee policies and procedures.

Chairperson: David A. Wilson, Ernst & Young
Jane O. Burns, Texas Tech University
Robert J. Capettini, San Diego State University
Konrad W. Kubin, Virginia Polytechnic Institute and State University
Joseph J. Schultz, Arizona State University
Paula B. Thomas, Middle Tennessee State University

FINANCIAL ACCOUNTING STANDARDS COMMITTEE

Charges: 1. To evaluate selected discussion memoranda and exposure drafts related to financial accounting and reporting in the private sector as they are released by the FASB and to respond to the FASB in writing and by appearing at selected public hearings.

2. To be cognizant of emerging issues related to financial accounting and reporting in the private sector through materials issued by the Emerging Issues Task Force and other relevant sources with evaluation, where appropriate, of the significance and potential implications of such issues.

3. To meet with the FASB, normally on an annual basis.

4. To explore the publication of selected responses by the committee in *Horizons* or other appropriate outlets.

5. To consider and promote ways to increase the academic input to the standard-setting process.

6. To provide input to the Financial Reporting Issues Conference Committee on the appropriate content and structure of the 1992 conference.

Chairperson: Jerry J. Weygandt, University of Wisconsin-Madison
Mary E. Barth, Harvard University
William A. Collins, University of North Carolina at Greensboro
G. Michael Crooch, Arthur Andersen & Co.
Thomas J. Frecka, University of Notre Dame
Eugene A. Imhoff, Jr. University of Michigan
Charles L. McDonald, University of Florida
Lawrence Revsine, Northwestern University
D. Gerald Searfoss, Deloitte & Touche

FINANCIAL REPORTING ISSUES CONFERENCE COMMITTEE

Charges: 1. To plan the 1992 Financial Reporting Issues Conference in accordance with established AAA policies.

2. To seek the advice of the Financial Accounting Standards Committee on the content and structure of the 1992 conference.

3. To arrange to have the conference conducted and administered.

4. To report on the results and effectiveness of the conference and to recommend changes by November 1, 1992.

Chairperson: William W. Holder, University of Southern California
Loyd C. Heath, University of Washington
Maureen F. McNichols, Stanford University
D. Gerald Searfoss, Deloitte & Touche
Lawrence A. Tomassini, University of Illinois

INNOVATION IN ACCOUNTING EDUCATION AWARD COMMITTEE

Charge: To select recipient(s) of the Outstanding Teaching/Curriculum Development Award using the criteria approved by the Executive Committee.

Chairperson: George W. Krull, Jr., Grant Thornton
Severin V. Grabski, Michigan State University
Jeffrey L. Harkins, University of Idaho
Marc F. Massoud, Claremont McKenna College
Karen V. Pincus, University of Southern California
Jacky Sanders, Mercer County Community College
Hadley P. Schaefer, University of Florida

INTERNATIONAL FACULTY EXCHANGE COMMITTEE

Charges: 1. To select the professors to be involved in the program for 1992–1993.

2. To assist the Executive Director with announcements, arrangements, and follow-up evaluations for the respective visits.

Chairperson: Trevor S. Harris, Columbia University
Sidney J. Gray, Glasgow University
Paul M. Healy, Massachusetts Institute of Technology
Gary K. Meek, Oklahoma State University
Mohamed Onsi, Syracuse University
Juan M. Rivera, University of Notre Dame

MINORITY FACULTY DEVELOPMENT COMMITTEE

Charges: 1. To summarize the recommendations made by the Minority Faculty Development Committees for the past five years and evaluate the impact these recommendations are having on the Association's activities.

2. To explore additional ways in which potential minority candidates can be encouraged to apply to accounting doctoral programs and to pursue professional careers in accounting education.

3. To explore ways in which to encourage the development of minority faculty members.

4. To examine and monitor minority recruitment programs of other professional organizations, such as the AICPA.

Chairperson: Joyce N. Holley, Texas A & M University
V. Gopalakrishnan, George Mason University
Margaret Hicks, Howard University
Robert R. Oliva, Florida International University
Jenice P. Stewart, Olin School of Business

NEW FACULTY CONSORTIUM COMMITTEE

Charges: 1. To plan the 1992 New Faculty Consortium in accordance with established AAA policies.

2. To arrange to have the consortium conducted and administered.

3. To report on the results and effectiveness of the consortium and to recommend changes by March 1, 1992.

Chairperson: Michael D. Diamond, University of Southern California
Silvia A. Madeo, University of Missouri-St. Louis
Dennis Reigle, Arthur Andersen & Co.
Mikel G. Tiller, Indiana University

NOMINATIONS COMMITTEE

Charge: To select a list of nominees to AAA offices for election by the membership in August 1992.

Chairperson: Gerhard G. Mueller, University of Washington
Alvin A. Arens, Michigan State University
Kathleen R. Bindon, University of Alabama
William L. Felix, Jr., University of Arizona
Jeffrey L. Harkins, University of Idaho
John K. Simmons, University of Florida
John J. Willingham, KPMG Peat Marwick

NOTABLE CONTRIBUTIONS TO ACCOUNTING LITERATURE AWARD SCREENING COMMITTEE

Charge: To identify works of exceptional merit from published accounting books and articles that meet the guidelines for the "Notable Contributions" awards established by the Executive Committee.

Chairperson: John H. Evans III, University of Pittsburgh
Randolph P. Beatty, University of Chicago
Bruce Bublitz, University of Kansas
Owen J. Cherrington, Brigham Young University
John Cumming, Miami University
Ronald A. Dye, Northwestern University
William L. Ferrara, Stetson University
Dale L. Flesher, University of Mississippi
Darrel I. Gosse, Indiana University
G. William Glezen, University of Arkansas
Michael H. Granof, University of Texas at Austin
James A. Heintz, University of Connecticut
Harvey S. Hendrickson, Florida International University
Orville R. Keister, University of Akron
Larry N. Killough, Virginia Polytechnic Institute and State University
Kathryn M. Means, Florida Atlantic University
Kathleen T. McGahran, Columbia University
James C. McKeown, Penn State University
Hugo Nurnberg, CUNY-Baruch College
Gordian A. Ndubizu, Drexel University
Sue Ravenscroft, Eastern Michigan University
Jane L. Reimers, Florida State University
Ashwinpaul C. Sondhi, New York University
Clyde P. Stickney, Dartmouth College
Murat N. Tanju, University of Alabama at Birmingham
Richard L. Townsend, University of Tennessee
Charles A. Tritschler, Purdue University
Richard G. Vangermeersch, University of Rhode Island
G. Peter Wilson, Harvard University
Rhoda C. Icerman, Florida State University

NOTABLE CONTRIBUTIONS TO ACCOUNTING LITERATURE AWARD SELECTION COMMITTEE
(JOINT COMMITTEE WITH THE AICPA)

Charge: To select the recipient(s) of the award from those books and articles identified by the Screening Committee as potentially notable accounting literature contributions.

Chairperson: Carl S. Warren, University of Georgia
Dan S. Dhaliwal, University of Arizona
Edward J. Schnee, University of Alabama
Wayne H. Shaw, Cornell University

OUTSTANDING ACCOUNTING EDUCATOR AWARD COMMITTEE

Charge: To select one or, at most, two recipients of the Outstanding Educator Award using the criteria approved by the Executive Committee.

Chairperson: Wanda A. Wallace, College of William and Mary
Andrew D. Bailey, Jr., University of Arizona
William A. Hillison, Florida State University
Gerald L. Salamon, Indiana University
E. Dan Smith, University of Georgia
Jean C. Wyer, Coopers & Lybrand

PROFESSIONAL EXAMINATIONS COMMITTEE

Charges: 1. To monitor activities of the AICPA and NASBA and evaluate the implications for the CPA examination and accounting education.

2. To identify activities the Association should consider undertaking that deal with the effect of professional examinations on accounting education.

3. To coordinate all activities and recommendations with the Accounting Education Advisory Committee.

Chairperson: Alan Reinstein, Wayne State University
James Blum, American Institute of CPAs
O. Whitfield Broome, University of Virginia
Patrick R. Delaney, Northern Illinois University
John M. Hassell, University of Texas at Arlington
Anthony T. Krzystofik, University of Massachusetts
Kenneth A. Smith, Idaho State University

PROFESSIONALISM AND ETHICS SEMINAR COMMITTEE

Charges: 1. To plan the 1992 Professionalism and Ethics Seminar in accordance with established AAA policies.

2. To arrange to have the seminar conducted and administered.

3. To continue the development of suitable educational materials related to professionalism and ethics.

4. To report on the results and effectiveness of the seminar and to recommend changes by July 1, 1992.

Chairperson: Stephen E. Loeb, University of Maryland
Mary Beth Armstrong, California Polytechnic State University
James C. Gaa, McMaster University
William May, University of Southern California
Joanne Rockness, North Carolina State University
Robert G. Ruland, Northeastern University
Robert J. Sack, University of Virginia

PROGRAM ADVISORY COMMITTEE

Charge: To assist the President and the Executive Director in developing the technical program for the 1992 Annual Meeting.

Chairperson: Frank R. Rayburn, University of Alabama at Birmingham
Vice-Chairperson: Mary S. Stone, University of Alabama
Area Coordinator: L. Howard Godfrey, North Carolina at Charlotte
Vice-Chairperson: Loren A. Nikolai, University of Missouri at Columbia
Area Coordinator: Jack L. Krogstad, Creighton University
Area Coordinator: E. Kent St. Pierre, James Madison University
Richard E. Baker, Northern Illinois University
J. Richard Dietrich, University of Texas at Austin
Michael L. Ettredge, Washington State University
Gerald A. Feltham, University of British Columbia
Lynford E. Graham, Coopers & Lybrand
Donald J. Kirk, Columbia University
John L. Kramer, University of Florida
Barry L. Lewis, University of Colorado at Boulder
Grant W. Newton, Seaver College
Edmund Outslay, Michigan State University
James E. Rebele, Lehigh University
Katherine Schipper, University of Chicago
David E. Stout, Villanova University
Mikel Tiller, Indiana University

Sections Liaison Representatives

Accounting Behavior and Organizations Section—Frank Collins, University of Miami
Administrators of Accounting Programs Group—Michael A. Diamond, University of Southern California
American Taxation Association—Shirley Dennis-Escoffier, University of Miami
Auditing—Jean Bedard, University of Pittsburgh
Artificial Intelligence and Experts Systems Section—Miklos A. Vasarhelyi, Rutgers University-Newark
Gender Issues in Accounting (Provisional Section)—Camille Koehler, University of New Mexico
Information Systems/Management Advisory Services—Rebecca Reuber, University of Toronto
International Accounting Section—Timothy S. Doupnik, University of South Carolina
Management Accounting—Henry R. Anderson, University of Central Florida
Two-Year College—Patrick Reihing, Nassau Community College
Teaching & Curricula—Bill N. Schwartz, Virginia Commonwealth University

RESEARCH ADVISORY COMMITTEE

Charges: 1. To monitor the accounting research activities of other accounting associations, e.g., AICPA, FEI, IIA, and NAA.

2. To advise the Director of Research in the development and administration of the Association's program in accounting research.

3. To suggest ways to encourage research in accounting, especially applied research dealing with all accounting functions in enterprise settings.

Chairperson: Nicholas Dopuch, Washington University
Rick Antle, Yale University
Ronald A. Dye, Northwestern University
Abbie Smith, University of Chicago
Dan A. Simunic, University of British Columbia
Ira Solomon, University of Illinois

SECURITIES AND EXCHANGE COMMISSION LIAISON

Charges: 1. To prepare and transmit a response to the staff of the Securities and Exchange Commission on proposals soliciting views from the public concerning auditing and financial accounting and reporting.

2. To make recommendations and to develop program(s) that will assist teaching of and conducting research in topics concerning the SEC and its activities.

3. To conduct activities as appropriate to assist communications and interactions between the SEC and the membership of the Association.

Chairperson: Thomas R. Weirich, Central Michigan University
Charles W. Mulford, Georgia Institute of Technology
Robert W. Rouse, Auburn University
Robert H. Sanborn, University of Richmond
David B. Smith, Claremont McKenna College

SENIOR FACULTY CONSORTIUM COMMITTEE

Charges: 1. To plan the 1992 Senior Faculty Consortium in accordance with established AAA policies.

2. To arrange to have the consortium conducted and administered.

3. To report on the results and effectiveness of the consortium and to recommend changes by August 1, 1992.

Chairperson: Helen Gernon, University of Oregon
Stephen T. Limberg, University of Texas at Austin
Katherine Schipper, University of Chicago
Gerald Smith, University of Northern Iowa

SEVENTH INTERNATIONAL CONFERENCE ON ACCOUNTING EDUCATION (1992) PLANNING COMMITTEE

Charge: To continue planning for the conference.

Chairman: Belverd E. Needles, Jr., DePaul University
Thomas A. Gavin, University of Tennessee at Chattanooga
Sidney J. Gray, Glasgow University
W. Morley Lemon, University of Waterloo
William Markell, University of Delaware
Seigo Nakajima, Ferris Jogakuin Women's College
Walter F. O'Connor, Fordham University
Norlin G. Rueschoff, University of Notre Dame
Murray C. Wells, University of Sydney

TRUEBLOOD SEMINARS COMMITTEE

Charges: 1. To assist Deloitte & Touche in the conduct of the Trueblood Seminars.

2. To report on the results and effectiveness of the seminars and to recommend changes by August 1, 1992.

Chairperson: R. D. Nair, University of Wisconsin-Madison
Sarah York Kenny, University of Utah
James G. Manegold, University of Southern California
Jack E. Wilkerson, Jr., Wake Forest University

THE ACCOUNTING REVIEW
Vol. 67, No. 1
January 1992
pp. 230–239

BOOK REVIEWS

Shane R. Moriarity, Editor

Editor's Note: Two copies of books for review should be sent to Professor Shane Moriarity, School of Accounting, University of Oklahoma, Norman, OK 73019. The policy of the *Review* is to publish only those reviews solicited by the Book Review Editor. Unsolicited reviews will not be accepted.

JOHN HAWLEY ATKINSON, JR., GREGORY HOHNER, BARRY MUNDT, RICHARD B. TROXEL, and WILLIAM WINCHELL, *Current Trends in Cost of Quality: Linking the Cost of Quality and Continuous Improvement* (Montvale, New Jersey: National Association of Accountants, 1991, pp. xi, 114, $34.95, paper).

This monograph is a product of the Institute of Management Accountants' (IMA) (formerly the National Association of Accountants) Bold Step Research Series. The purpose of this series is to "improve management accounting by identifying state-of-the-art techniques and showing how these techniques can be adapted to changing manufacturing and service technologies." This project was proposed by KPMG Peat Marwick, which employs three of the authors.

Increased competition in a globalized economy has made quality a significant feature in distinguishing between products and services. As a result, many corporations have implemented quality control and quality improvement programs, and monitor the progress of these programs with a quality cost system. The purpose of this monograph is to "identify key organization and business practices necessary to implement a cost of quality system and support its effective, ongoing operation." The monograph is designed to extend the work of a previous monograph, *Measuring, Planning and Controlling Quality Costs*, by Morse, Roth, and Poston (also published by the IMA) which introduces quality cost concepts and quality cost reporting systems.

The monograph is divided into three parts. The first two chapters introduce the concepts of quality costing. Chapters 3 through 6 are descriptions of case studies performed by the authors on four corporations. These corporations include Banc One, a multi-state bank holding company; Cascade Engineering, a custom injection molding manufacturer; the Direct Marketing Division of Hewlett-Packard Company; and the Supplies Development and Manufacturing Unit of Xerox Corporation. The remainder of the monograph (Chapters 7 through 10) consists of interpretation and implications of the case study findings.

The monograph is successful in achieving its purpose of identifying strategies for implementing and operating quality cost systems. The primary finding of the study relates to the use of quality cost systems. While quality cost systems have been viewed traditionally as a subset of the formal accounting system (as suggested by Morse, Roth, and Poston), the authors found that the quality cost system is more accurately represented as a separate, independent system. They also found quality cost reporting is an informal process (one which is not driven by rigid reporting periods) and has an operational (product- and process-oriented), rather than financial (cost reduction) focus (Chapter 7). In light of these findings, the authors suggest an updated model for implementing a quality cost system that accommodates the observed uses of these systems (Chapter 8).

In addition, the monograph has an excellent discussion of the role of the quality cost system within the operational philosophy of "continuous quality improvement" (Chapter 9). There is also a discussion of future trends in quality costing and the effects on quality improvement (Chapter 10).

The primary weakness of the monograph is in its description of the quality cost categories and the relationships between these categories (Chapter 2). A confusing description of quality cost relationships was introduced before defining quality costs. The authors propose a "modern view" of quality cost relationships with insufficient evidence or logic to support the view.

Excluding this weakness, the monograph is a commendable supplement to Morse, Roth, and Poston's previous work. The monograph would certainly be useful to those implementing or maintaining a quality cost system in their corporation. As many authors have not yet integrated detailed quality cost information in cost accounting texts, academicians may find the monograph a useful source for teaching quality costs in managerial accounting.

WILLIAM R. PASEWARK
Assistant Professor
University of Houston

BHABATOSH BANERJEE, editor, *Contemporary Issues in Accounting Research* (Calcutta: Indian Accounting Research Foundation, 1991, pp. xvi, 338, $25.00).

The Indian Accounting Association (IAA) was formed in 1969 in an effort to consolidate the profession of accountancy in India along the lines of the American Accounting Association. In 1990, the IAA established the Indian Accounting Association Research Foundation (IAARF) to promote fundamental and applied research in accounting and other allied areas in India. *Contemporary Issues in Accounting Research* is the first publication of this newly established body. This research volume is edited by Professor Bhabatosh Banerjee, President of the Indian Accounting Association and former Dean of Faculty of Commerce, Social Welfare and Business Management at University of Calcutta, India. The stated objective of this research volume is "to provide a base for lively dialogue on a number of serious issues concerning accounting research and to serve as the gateway to further deliberations and research on the subject" (p. ix).

The volume consists of 18 unrelated invited research papers by authors from the United States, the United Kingdom, Canada, India, and Japan. As with any collection of research papers, some are more insightful than others. From my perspective, the opening paper, "Social versus Physical Reality in Accounting and the Measurement of its Phenomena," by Professor Richard Mattessich is the most interesting in the entire volume. Of the remaining 17 papers, four are in the area of international accounting, five highlight issues pertinent to India, while four review the existing literature with little or no future directions for research. The remaining four papers deal with such areas as a cash flow approach to accounting, JIT, the role of accounting in varied socio-political environments, and social responsibility in accounting.

Interestingly, the editor has completely ignored the area of auditing. Editors who are faced with compiling the first edition of such a work must make difficult choices with respect to topics and depth of coverage. However, for completeness, I would have preferred to see a few papers in auditing.

The individual papers differ greatly in quality. A number of them are verbose and not well organized. In some, the reader is left guessing as to what the author will discuss next or even what the purpose of the paper may be. There is also a lack of consistency in format and style. For example, some papers provide an introduction and a conclusion, while others skip either one or both. Some papers use reference citations that follow the usual U.S. style while others use endnotes. In addition, the papers in this volume are arranged in a random order. For example, a paper treating social versus physical reality is followed by a review of the literature in positive accounting theory which is then followed by Professor Kenneth Most's paper suggesting avenues of future research in international accounting. This paper is followed by another paper that reviews accounting information and earnings prediction research, and so on. This random order may make the research volume unattractive for classroom use in North America. The volume could have been much leaner and more readable if more careful editing and planning had been used.

Pedagogically, this research volume will serve better as a textbook in India and nearby countries than in the United States. In the developing countries of Asia this research volume could be used for supplementary reading, both at the master's and the Ph.D. level. In addition, some of the research papers highlighting accounting problems in developing countries could be used in a graduate level international accounting course in Western countries.

Despite the shortcomings of the research volume, I am impressed with the effort put forth by the Indian Accounting Association Research Foundation (IAARF). As far as I am aware, this is the first English-language book on accounting research from a developing country. In summary, the IAARF has taken a first step that is long overdue. It has begun to disseminate scholarly research in India.

PARVEEN P. GUPTA
Assistant Professor of Accounting

Lehigh University

JOHN INNES and FALCONER MITCHELL, *Activity Based Costing: A Review with Case Studies* (London: The Chartered Institute of Management Accountants, 1990, pp. x, 70, £15.95, paper).

This book, commissioned by the Chartered Institute of Management Accountants, is intended to serve the growing practitioner interest in Activity Based Costing (ABC) in the U.K. Although the book draws heavily on the work of American academics, it nevertheless strives to present the material in a manner accessible to U.K. practitioners. The book explains the principles and mechanics of ABC in a clear and cogent manner. The authors provide numerous diagrams, worked examples, and three U.K. based case studies.

The authors compare and contrast the allocation of overheads using such cost drivers as material movements, number of orders, number of inspections, etc., with more conventional volume based drivers such as direct labor hours, to reveal considerable discrepancies between the two methods. Whilst illustrative of the method, worked examples of this kind always lend themselves to the charge that they were designed to reveal and possibly exaggerate such discrepancies. In consequence they may oversell the potential of the new technique.

This charge is recognized by the authors in the second section of the book which provides "a critical review of ABC." It is this section which gives the book its distinguishing features. On a positive note, the critical review cautions us against becoming too fascinated with the technique of ABC. The authors suggest that we should consider ABC within the wider context of its contribution to the management and decision making process of both service and manufacturing organizations. The problem with this section, and the one that serves to disappoint the reader, is that although it raises a number of important issues on the scope, purpose, differentiating features, procedural difficulties and behavioral effects of ABC, it falls far short of a truly critical analysis.

For example, the authors note that ABC has a similar framework to that of a conventional costing system. "First, costs are pooled and second, a series of cost rates are used to attach the pooled costs to products" (20). However, rather than pursuing this critical line of enquiry which could raise valid questions about whether ABC constitutes a new "method" of costing or merely a new "technique" of calculating cost rates, the authors simply report, in a largely uncritical manner, the differentiating features of ABC offered by "advocates" of the system. This is also true of the sections on purpose and decision orientation. Here the authors again begin to recognize some of the limitations of ABC but all too readily capitulate to the lines of argument offered by leading advocates of the technique.

The book further compensates for some of the deficiencies of using worked examples by providing three short case studies. Although the cases, like the preceding discussion, are limited in terms of their critical content, they nevertheless begin to touch upon the use of ABC in organizations by examining its impact in areas such as cost cutting, pricing, product line selection and the monitoring of overhead costs. As the section on future research suggests however, a considerable amount of work on the use, impact, and scope of the ABC remains to be done.

Despite the limitations of this work, it provides U.K. based practitioners with a good overview of the techniques of ABC and addresses some its potential benefits and limitations. The book is also a valuable teaching resource and serves as a useful bibliography of the ABC literature.

ALISTAIR M. PRESTON
Assistant Professor of Accounting

Boston University

MARC P. JOYE and PAUL J. BLAYNEY, *Cost and Management Accounting Practices in Australian Manufacturing Companies: Survey Results* (Sydney: Accounting and Finance Foundation, University of Sydney, Monograph No. 7, 1990, pp. xiii, 219, A$25.00, paper).

This book can be characterized as a technical report—an informative summary and elaboration upon an extensive data base on cost and management accounting practices. As such, the book does not devote extensive energy to the development of theoretical frameworks or related analysis of results. Even the authors acknowledge that comments and analysis are included in related working papers published elsewhere. Instead, the strength of the book is the presentation of some preliminary and comprehensive descriptive data pertaining to several critical issues raised in contemporary management accounting literature. This report is essentially motivated by the fact that little empirical data exists to support the current literature which suggests that cost accounting is in crisis and is the basis for failure in Western manufacturing companies.

The wealth of observations generated in this report results from a survey of over 2,000 large manufacturing firms in Australia, of which 430 responded. These respondents were primarily from four industries—food, chemical product, fabricated metal, and machines and computers. The data gathering instrument itself is presented and management accounting researchers should find this 31-item questionnaire informative and enlightening. This extensive questionnaire covers such issues as diversity of the firm (products, line units, cost centers, profit centers), behavioral issues (performance evaluation, capital budget requests, incentive compensation, setting objectives, decentralization of authority and decision-making), cost accounting practices (allocation of overhead, product costing systems, variable versus fixed cost definitions), and pricing and internal transfer pricing issues.

The data presents preliminary evidence to test some of the speculations and assumptions which are being raised in current managerial accounting literature. For example, the findings suggest that product pricing seems to be the main reason for overhead cost allocation, that product costs are the primary basis for pricing (as compared to marketing and competitive strategies), and that volume-based allocations (direct labor, units of production, machine hours) are still predominantly used. This set of related findings motivate the authors to raise the following provocative issue: Given this reliance on the cost accounting system, this doesn't seem to bode well for Australian firms in light of the recent literature criticizing cost accounting techniques. On this point, however, the authors state that the challenge to researchers is to demonstrate whether cost-drivers, while conceptually attractive, pass the cost-benefit test because, as the authors state, to date the evidence is inconclusive.

The report also raises interesting macro organizational and industry issues and the manner in which these variables might impact cost and management accounting systems. For example, as firms diversify, they tend to have more reasons for using multiple overhead rates. Furthermore, variations exist across industries in terms of the use of single versus multiple overhead rates. The report also observes that cost-based methods are primarily used for transfer pricing (as compared to market prices or negotiations). However, the more diversified the firm, the more market or negotiated prices are used; the less diversified, the more cost-based is used. These descriptive differences in the uses of cost and management accounting systems as a function of organizational and industry characteristics suggests the need for meaningful theoretical perspectives from which to examine managerial accounting practices.

In summary, this exploratory descriptive research does indeed provide a basis for practitioners and academics to increase their knowledge about manufacturing accounting practices and shed light for future research. This report serves as a helpful precursor and/or complement for those engaging in field and case studies on cost and management accounting practices.

MARK A. COVALESKI
Professor of Accounting

University of Wisconsin–Madison

RICHARD MACVE and JONATHAN JACKSON, 'Marking to Market': Accounting for Marketable Securities in the Financial Services Industry (London: The Institute of Chartered Accountants in England and Wales, 1991, pp. 155, £12.00, paper).

This study examines issues relating to valuing marketable securities at market with particular emphasis on institutions that deal in investments. The study was undertaken at the request of the Accounting Standards Committee of the U.K., and a substantial part of the study focuses on the relationship between marking to market, U.K. companies legislation, and accounting standards. However, given the increased attention from the AICPA, the SEC, and the FASB to market valuation of financial instruments in general, and debt securities in particular, this study should be of interest to a wider audience.

The approach taken is to first review the major conceptual and practical issues relating to accounting for investments and to reconcile accounting theory and practice. The authors do not draw a distinction between different types of investments and their analysis addresses the whole spectrum of investments including debt and equity securities, futures contracts, options, swaps, and warrants. The authors review the requirements of U.K. company laws, accounting standards, and of regulatory bodies as well as recommended practice in Japan, the U.S., and of the IASC. The study also includes interviews with auditors, regulators, and market makers as well as a survey of the practices reported in the financial statements of all the market makers recognized by the London Stock Exchange.

Some of the major conclusions of the study include:

- Knowledge of the market value of all investments is useful to users of financial statements. The relevant concept of valuation is "deprival value," which would normally be the bid price.
- However, economic analysis of the concept of income does not help to answer the issue of whether the change in market value should be included in income. It can only be answered by reference to the preferences and intentions of the holder of the investments.
- Since these preferences cannot be known, certain conventions must be chosen to operationalize the concept of income in accounting practice. These conventions would not treat normal fluctuations in the value of long-term investments as income but would count as income the fluctuations in value of most current investments and in particular of all dealing (i.e., trading) investments.
- In determining intentions, the organizational structure of the investment firm, i.e., trading versus treasury management versus long-term investment, should be evaluated.
- While the valuation basis should always be market, as a practical matter if the cost figure is more reliable then the lower of cost or market basis could be used. This would be a better approach than attempting to define situations where marking to market is or is not appropriate. Such definitions inevitably require arbitrary cut-offs.
- With hedges, one of the advantages of marking to market accounting is the simplicity achieved when both sides of the hedge can be marked to market.
- In countries where practice is developed primarily through accounting standards there is a movement towards accounting for current investments at current value. Where practice is governed primarily by a legal or quasi-legal code the practice of lower of cost or market is still predominant.
- As a general conclusion, the authors point out that a conceptual analysis of the problem does not help in addressing how to account for marketable investments. This failure, in their view, explains why rules devised by regulators for their specific purposes appear to drive the reporting of investments and the measurement of investment returns and income.

The authors provide a well-written discussion of the issues. The conceptual analysis is supplemented with numerous illustrations; the discussion of regulatory requirements is thorough; and the findings from the interviews and financial statement surveys are integrated into the text. It should be noted that in the U.S., the AICPA's proposal to account for debt securities on the basis of "management intent," a conclusion which the authors also reach, was dismissed by the SEC as "psychoanalytic accounting." Also, the authors' conclusions about the regulators strong influence on the "mark to market" debate is consistent with recent events in the U.S.

The study includes nine Appendices which address issues such as practice in the U.S. and Japan, guidance from the IASC, and a review of the financial statements of market makers on the London Stock Exchange. The review of practice in the U.S. focuses largely on the requirements of SFAS Nos. 52 and 80. Since the study was completed in October 1989, it does not address the recent FASB Exposure Draft on Disclosure of Market Values of Financial Instruments or the AICPA's attempts to address the market valuation of debt securities. The review of financial statements provides summary information on issues such as balance sheet valuation, income statement treatment, disclosures, and audit reports. Most companies used mark to market accounting in order to present a "true and fair" view of the balance sheet.

The book should be useful to all those interested in the debate over market valuation of financial instruments. Its blend of conceptual analysis, regulatory requirements, interviews and survey of financial statements provides perspectives and insights of interest to regulators, practitioners, and academics. Because of its description of practice and regulatory requirements in the U.K., this book could be used as a supplement for graduate courses in accounting theory or international accounting.

R. D. NAIR
Professor of Accounting
University of Wisconsin–Madison

JOSHUA RONEN, ANTHONY SAUNDERS, and ASHWINPAUL C. SONDHI, editors, *Off-Balance Sheet Activities* (Westport, CT: Quorum Books, 1990, pp. vii, 182, $39.95).

The dollar volume of off-balance sheet (OBS) banking activities has grown markedly in recent years. These transactions, which include loan commitments, future and forward contracts, standby letters of credit, options, swaps, and loan sales, comprise approximately 85 percent of the on-balance sheet assets of banks reporting on Schedule L to the Federal Reserve. For some money-center banks, OBS activities can run as high as four times a bank's balance-sheet reported activities. In recognition of the growth of OBS activities, the Federal Reserve is implementing new risk-based capital standards that incorporate the risk sof OBS activities into capital adequacy measures. These new capital standards will be effective at the end of 1992.

Responding to the growth and absolute economic significance of OBS activities, a New York University Conference was undertaken by the Vincent C. Ross Institute of Accounting Research and the Salomon Brothers Center for the Study of Financial Institutions at the Stern School of Business. The purpose of this conference was to bring together academic researchers in accounting and finance along with practitioners and regulators with the objective of gaining insights into, and proposing meaningful solutions to, the issues raised by the proliferation of OBS activities. The current book provides a hard-cover collection of the papers presented and discussants' comments rendered during that conference, divided into two main sections: Off-Balance Sheet Activities and Banking and Financial Reporting and Off-Balance Sheet Transactions.

Three research papers are presented in the first section. In "Off-Balance Sheet Activities and the Underinvestment Problem in Banking," Christopher James argues that loan sales and loans backed by standby letters of credit are an indirect means of securitizing deposits and can lead banks to make loans with positive net present values that they would pass up if restricted to unsecured, deposit financing. In "Are Loan Sales Really Off-Balance Sheet?" Gary Gorton and George Pennacchi find empirical evidence supporting their premise that loan sales are not really off-balance sheet. Loan sales contracts were found to contain risk premia implying the existence of implicit guarantees on the part of selling banks. In their paper, "Money Demand and Off-Balance Sheet Liquidity: Empirical Analysis and Implications for Monetary Policy," Reuven Glick and Steven Plaut find evidence to support their notion that firms and individuals demand liquidity in the form of both on-balance sheet and off-balance sheet instruments. This observation has policy implications because monetary targets are currently set using on-balance sheet instruments only. All three papers in this section were well-conceived and executed. Collectively, they help us to better understand why OBS activities exist and their implications for the risk of banking organizations and for money supply policy.

The second section contains four papers. Robert Swieringa's paper, "Recognition and Measurement Issues in Accounting for Securitized Assets," outlines the scope of the FASB's financial instruments project and focuses on recognition and measurement issues surrounding securitized assets. In "Contract Theoretic Analysis of Off-Balance Sheet Financing," Jim Donegan and Shyam Sunder conclude that contract enforcement considerations and an expansion of the range of financial reporting alternatives to include expected value reporting on one hand

and the reporting of contract sets on the other would facilitate the creation of intertemporally stable financial reporting standards. The paper, "The Use of Off-Balance Sheet Financing to Circumvent Financial Covenant Restrictions," by Samir El-Gazzar, Steven Lilien, and Victor Pastena presents evidence to support the premise that OBS activities are used to modify covenant-based restrictions on dividend distributions and the issuance of additional debt. In "Debt Capacity and Financial Contracting: Finance Subsidiaries," Joshua Ronen and Ashwinpaul Sondhi provide support for the point that captive finance companies were not created to hide the true nature of the debt-equity ratio, but rather to provide optimum financial benefit to the shareholders of the parent. Collectively, this set of papers provides a valuable discussion of issues impacting upon the recognition and measurement of OBS activities.

Researchers in accounting and finance, Ph.D. students, and certain advanced-level master's students will find "Off-Balance Sheet Activities" to be a useful book. The book is very timely and provides a good snapshot of where we are in the area of off-balance sheet transactions and their recognition and measurement. And because the authors are careful in defining terms and elaborating research questions, the book will have appeal not only to those who are knowledgeable in the area, but also to those who wish to become so. Finally, the appendix, which demonstrates how to compute capital ratios under the Federal Reserve's new risk-based capital guidelines, will be helpful to anyone interested in understanding the financial position of bank holding companies.

CHARLES W. MULFORD
Associate Professor of Accounting

Georgia Institute of Technology

MARK STEVENS, *The Big Six: The Selling Out of America's Top Accounting Firms* (New York: Simon & Schuster, 1991, pp. 271, $19.95).

The Big Six is a compendium of six topics that purport to describe the six largest accounting firms in the country and how they got to their current state of affairs. The six topics included are selective but are of considerable interest to members of the auditing profession and those interested in it. They include the ZZZZ Best case, Beverly Hills Savings and Loan, Arthur Andersen's splitting into two components, the evolution of Deloitte, Haskins and Sells, the mergers of Deloitte and Touche and Ernst and Young, and some insights in the proposal process central to competing for new clients. The author describes this latter process from his experience as a member of the audit committees of a couple of companies. The book is newsy with a lot of personal insight, quotes from individual players in the actions discussed and appears to represent a large amount of work combining interviews, news clippings, public court documents, and other sources of information. The text raises important questions about the current structure of auditing in this country and the major changes that are taking place.

On a less complimentary note, the overleaf or cover of the book claims the author is an expert on auditors in this country. This may or may not be true, but if Mr. Stevens is an expert, he tends to use the language of the plaintiff's bar and does not appear to have a great deal of interest in, knowledge of, or sympathy with the standards and the traditions of the profession. The language is often very flowery for academic taste and leaves one concerned about separating fact from opinion. A careful reader will also note that many statements are not supported by evidence. Perhaps this is understandable in a book that does not claim to be scholarly. Also, for my tastes, the two litigation cases selected, ZZZZ Best and Beverly Hills Savings, are not the most interesting among several possible choices. It may have been the best available at the time the author was writing the text, however. Reading these two chapters strongly suggests a need for a more thorough analysis of recent litigation experience, including civil, criminal, and SEC 2E proceedings. As an academic, I would like to look at a book like this with a view towards using it in the classroom. However, the language of the profession is used inaccurately or incorrectly with enough frequency that I would not recommend it. To illustrate this point, on page 43 in discussing SAS 16, the following phrase appears: "errors or irregularities [euphemisms for 'fraud']." This series of words, of course, suggests that the author does not know the difference between errors and irregularities or does not know the definition of fraud. On page 80 in discussing reporting, the following is written:

> When an S&L is blessed with an unqualified opinion complete with the prestigious signature of a prominent CPA firm, the public rests assured that it can take *comfort* in its financial statements. [italic added]

This sentence strongly suggests the author believes that auditors are offering fiscal assurance rather than disclosure assurance. Another indication of this view is on page 64 where the author states, "... have gone belly-up only weeks or months after the auditors assured the world of their fiscal well-being." The message in financial statements may be anything but comforting, but the author suggests strongly that a clean opinion indicates otherwise.

For the more experienced student of auditing, this book has some useful messages. It puts in one place a lot of news regarding recent events in the profession. It describes rather clearly how outsiders' reasonable perceptions of some of these recent events can be extremely negative. Such a view can go a long way towards explaining how some users and some congressmen obtain the perceptions they have of the profession.

The Big Six is an enjoyable book for more experienced students. From the view of the plaintiff's bar, it is not a perfect indictment, but it has the right style of language. For a person who is more sympathetic to the profession and its traditions, the language is far too strong, but there are still important lessons to be learned. In summary, I would recommend this text for academics and possibly more advanced classes, but certainly not for introductory auditing classes.

WILLIAM L. FELIX, JR.
Professor of Accounting

University of Arizona

PAULINE WHITE and CONSUELO M. OHANESIAN, *1991 Tax Penalties: The Complete Guide to Penalties Under the Internal Revenue Code* (Naperville, IL: Financial Sourcebooks, 1991, pp. xiii, 362, $49.95, paper).

This book, written by corporate attorneys, consists of 75 chapters organized by specific tax code sections related to tax penalties, ranging from failure to file a return ... to paying your tax with a bad check ... to refusal to permit entry or examination by an officer or employee of the Treasury Department. Each chapter includes two parts, the first quoting the full text of the I.R.S. statute and the second, which is entitled "Interpretation," rehashing the statute in the authors' language plus a list of questions, answers, and definitions. It is perhaps this last section which possesses the greatest potential for usefulness as a resource tool, for here the authors refer to interpretative case law, rulings, regulations, and legislative history, albeit briefly. The authors' stated objective in preparing the volume was that: "first, our readers could avoid penalties by understanding what conduct was being penalized; and second, to aid in establishing that a penalty's application was incorrect" (p. x).

While I found the stated objective intriguing and the cases cited worthwhile, I felt that the book would benefit greatly from an expanded index. The index is only slightly over three pages long, in contrast to the table of contents, which is nearly six pages long. For example, neither exempt organizations nor rollover distributions are listed in the index; both of these topics are covered in Chapter 2, which is entitled merely "Failure to File Certain Information Returns." Rather than being organized by topic, the table of contents and the book itself are arranged from lowest I.R.S. section number to highest, which would be a reasonable approach if the index were more thorough.

The authors claim to describe suggested defenses for combating penalties; and while these defenses are relatively rare in the 362 page book, they are among the more valuable tidbits. For example, reasonable cause is mentioned as a plausible defense in some cases and is defined in rather technical terms related to exercising "ordinary business care and prudence" and the inability to pay or file or "undue hardship" (p. 8). However, this somewhat vague definition is followed by a specific list including, among others, death or serious illness of the taxpayer or a member of his immediate family, destruction of the taxpayer's residence, place of business, or business records by fire or other casualty, and the receipt of erroneous information from an I.R.S. employee.

This is the only book of its kind (devoted solely to describing tax penalties, both civil and criminal and how they are determined and assessed). Tax penalties can amount to big dollars, and if this guidebook can serve to point the way for an accountant or lawyer to help taxpayers avoid penalties, then the purchase of the book would be justified. It might be noted, though, that this is a 1991 edition, which suggests the necessity to repeat the expenditure annually if one is to be sure of staying current. Also, the authors have excluded law applicable to years prior to 1991.

DEBRA C. JETER
Associate Professor of Accounting
Austin Peay State University

BOOKS RECEIVED MAY 8, 1991 TO AUGUST 10, 1991

JERRY W. ANDERSON, JR., and JOHN B. CAMEALY, *Corporate Operational Analysis: A Procedure for Evaluating Key Factors in Internal Operations, Acquisitions, and Takeovers* (New York: Quorum Books, 1991, pp. xiii, 347, $59.95).

JOHN ARNOLD, PAUL BOYLE, ANTHONY CAREY, MALCOLM COOPER, and KEN WILD, *The Future Shape of Financial Reports* (London: The Institute of Chartered Accountants in England and Wales and The Institute of Chartered Accountants of Scotland, 1991, pp. iv, 36, £9.95, paper).

JOHN HAWLEY ATKINSON, JR., GREGORY HOHNER, BARRY MUNDT, RICHARD B. TROXEL, and WILLIAM WINCHELL, *Current Trends in Cost of Quality: Linking the Cost of Quality and Continuous Improvement* (Montvale, NJ: National Association of Accountants, 1991, pp. xi, 114, $34.95, paper).

AHMED BELKAOUI, *Multinational Management Accounting* (New York: Quorum Books, 1991, pp. xii, 292, $59.95).

RAYMOND C. COLE, JR., and H. LEE HALES, *CIM Justification in the Process Industry: A Case Study* (Montvale, NJ: National Association of Accountants, 1991, pp. ix, 51, $12.50, paper).

COOPERS & LYBRAND, *1991 International Accounting Summaries: A Guide for Interpretation and Comparison* (New York: John Wiley & Sons, Inc., 1991, pp. xvi, 667, $95.00).

DONALD A. CURTIS, *Management Rediscovered: How Companies Can Escape the Numbers Trap* (Homewood, IL: Richard D. Irwin, Inc., 1990, pp. xi, 200, $24.95).

W. E. DIEWERT, *Contributions to Economic Analysis 196: Price Level Measurement* (Amsterdam: Elsevier Science Publishers B. V., 1990, pp. vii, 377, $69.50).

DALE L. FLESHER, *The Institute of Internal Auditors: 50 Years of Progress Through Sharing* (Altamonte Springs, FL: The Institute of Internal Auditors, 1991, pp. xi, 180, $15.00 members, $20.00 non-members).

C. G. A. FLETCHER, *Government Accounting* (Edinburgh: The Institute of Chartered Accountants of Scotland, 1991, pp. vii, 148, £7.50, paper).

JACK FOX, *Starting and Building Your Own Accounting Business*, 2nd ed. (New York: John Wiley & Sons, Inc., 1991, pp. xvii, 394, $49.95).

O. FINLEY GRAVES, ed., *The Costing Heritage: Studies in Honor of S. Paul Garner* (Harrisonburg, VA: The Academy of Accounting Historians, 1991, pp. ix, 171, $15.00 individuals, $20.00 institutions).

S. J. GRAY, C. B. ROBERTS, and P. D. GORDON, *Making Corporate Reports Valuable: A Survey of Corporate Reporting Practices by Major UK Companies* (Edinburgh: The Institute of Chartered Accountants of Scotland, 1991, pp. Vol. 1: 43, Vol. 2: 212, £20.00, paper).

DAVID J. HARR and JAMES T. GODFREY, *Private Sector Performance Measures and Their Applicability to Government Operations* (Montvale, NJ: National Association of Accountants, 1991, pp. xi, 93, $14.95).

TAKASHI KANATSU, *TQC for Accounting: A New Role in Companywide Improvement* (Cambridge: Productivity Press, English translation 1990, pp. xviii, 226, $45.00).

BRIAN H. MASKELL, *Performance Measurement for World Class Manufacturing: A Model for American Companies* (Cambridge: Productivity Press, Inc., 1991, pp. xxi, 408, $45.00).

ROY L. NERSESIAN, *Computer Simulation in Financial Risk Management: A Guide for Business Planners and Strategists* (New York: Quorum Books, 1991, pp. xiv, 224, $47.95).

BENNIE H. NUNNALLY, JR., D. ANTHONY PLATH, and HELENE W. JOHNS, *Corporate Lease Analysis: A Guide to Concepts and Evaluation* (New York: Quorum Books, 1991, pp. ix, 242, $49.95).

DOUGLAS A. POTTER, *Automated Accounting Systems and Procedures Handbook* (New York: John Wiley & Sons, Inc., 1991, pp. xii, 562, $105.00).

ESBJÖRN SEGELOD, *Capital Investment Appraisal: Towards A Contingency Theory* (Lund: Student litteratur, 1991, pp. 88).

HOUSTON H. STOKES, *Specifying and Diagnostically Testing Econometric Models* (New York: Quorum Books, 1991, pp. xiii, 337, $55.00).

ROBERT N. SUGHRUE and MICHELLE L. KOPNSKI, *Qualifying as a Nonprofit Tax-Exempt Organization: A Guide for Attorneys, Accountants, and Executive Management* (New York: Quorum Books, 1991, pp. vii, 220, $39.95).

G. A. SWANSON and HUGH L. MARSH, *Internal Auditing Theory—A Systems View* (New York: Quorum Books, 1991, pp. xv, 217, $47.95).

CHARLES J. TEPLITZ, *The Learning Curve Deskbook: A Reference Guide to Theory, Calculations, and Applications* (New York: Quorum Books, 1991, pp. xix, 288, $49.95).

WALTER P. VON WARTBURG, *How to Lead with Genius: Words of Wisdom for the Common Sense Manager* (New York: Markus Wiener Publishing, Inc. and Shapolsky Publishers, Inc., 1991, pp. xi, 124, $15.95).

R. S. OLUSEGUN WALLACE, JOHN M. SAMUELS, and RICHARD J. BRISTON, editors, *Research in Third World Accounting: Volume 1—1990* (London: JAI Press Ltd., 1991, pp. xxi, 299, £44.50).

WANDA A. WALLACE, *Accounting Research Methods: Do the Facts Speak for Themselves?* (Homewood, IL: Richard D. Irwin, 1991, pp. xvii, 103, $44.95).

PAULINE WHITE and CONSUELO OHANESIAN, *1991 Tax Penalties: The Complete Guide to Penalties Under the Internal Revenue Code* (Naperville, IL: Financial Sourcebooks, 1991, pp. xiii, 362, $49.95, paper).

BENJAMIN WRIGHT, *EDI and American Law: A Practical Guide* (Alexandria, VA: The Electronic Data Interchange Association, 1989, pp. xxii, 114, $69.95).

EDITORIAL POLICY AND STYLE INFORMATION

EDITORIAL POLICY

According to the policies set by the Publications Committee (which were endorsed by the Executive Committee and were published in the *Accounting Education News*, June 1987), *The Accounting Review* "should be viewed as the premier journal for publishing articles reporting the results of accounting research and explaining and illustrating related research methodology. The scope of acceptable articles should embrace any research methodology and any accounting-related subject, as long as the articles meet the standards established for publication in the journal.... No special sections should be necessary. The primary, but not exclusive, audience should be—as it is now—academicians, graduate students, and others interested in accounting research."

The primary criterion for publication in *The Accounting Review* is the significance of the contribution an article makes to the literature.

The efficiency and effectiveness of the editorial review process is critically dependent upon the actions of both authors submitting papers and the reviewers. Authors accept the responsibility of preparing research papers at a level suitable for evaluation by independent reviewers. Such preparation, therefore, should include subjecting the manuscript to critique by colleagues and others and revising it accordingly prior to submission. The review process is not to be used as a means of obtaining feedback at early stages of developing the research.

Reviewers and associate editors are responsible for providing critically constructive and prompt evaluations of submitted research papers based on the significance of their contribution and on the rigor of analysis and presentation. Associate editors also make editorial recommendations to the editor.

MANUSCRIPT PREPARATION AND STYLE

The Accounting Review's manuscript preparation guidelines follow (with a slight modification) the B-format of the *Chicago Manual of Style* (13th ed.; University of Chicago Press). Another helpful guide to usage and style is *The Elements of Style*, by William Strunk, Jr., and E. B. White (Macmillan). Spelling follows *Webster's International Dictionary*.

FORMAT

1. All manuscripts should be typed on one side of 8½ × 11" good quality paper and be double spaced, except for indented quotations.
2. Manuscripts should be as concise as the subject and research method permit, generally not to exceed 7,000 words.
3. Margins of at least one inch from top, bottom, and sides should facilitate editing and duplication.
4. To assure anonymous review, authors should not identify themselves directly or indirectly in their papers. Single authors should not use the editorial "we."
5. A cover page should show the title of the paper, the author's name, title, and affiliation, any acknowledgements, and a footnote indicating whether the author would be willing to share the data (see last paragraph in this statement).

Pagination: *All* pages, including tables, appendices, and references, should be serially numbered. The first section of the paper should be untitled and unnumbered. Major sections may be numbered in roman numerals. Subsections should not be numbered.

Numbers: Spell out numbers from one to ten, except when used in tables and lists, and when used with mathematical, statistical, scientific, or technical units and quantities, such as distances, weights, and measures. For example: *three days*; *3 kilometers*; *30 years*. All other numbers are expressed numerically. Generally when using approximate terms spell out the number, for example, *approximately thirty years*.

Percentages and Decimal Fractions: In nontechnical copy use the word *percent* in the text; in technical copy the symbol % is used. (See the *Chicago Manual* for discussion of these usages.)

Hyphens: Use a hyphen to join unit modifiers or to clarify usage. For example: a *well-presented analysis*; *re-form*. See *Webster's* for correct usage.

Key Words: The synopsis is to be followed by four *key words* that will assist in indexing the paper.

SYNOPSIS

A synopsis of about 400 words should be presented on a separate page immediately preceding the text. The synopsis also serves as the "introduction." The synopsis should be nonmathematical and include a readable summary of the research question, method, and the significance of the findings and contribution. The title, but not the author's name or other identification designations, should appear on the synopsis page.

TABLES AND FIGURES

The author should note the following general requirements:

1. Each table and figure (graphic) should appear on a separate page and should be placed at the end of the text. Each should bear an arabic number and a complete title indicating the exact contents of the table or figure.
2. A reference to each graphic should be made in the text.
3. The author should indicate by marginal notation where each graphic should be inserted in the text.
4. Graphics should be reasonably interpreted without reference to the text.
5. Source lines and notes should be included as necessary.

Equations: Equations should be numbered in parentheses flush with the right-hand margin.

DOCUMENTATION

Citations: Work cited should use the "author-date system" keyed to a list of works in the reference list (see below). Authors should make an effort to include the relevant page numbers in the cited works.

1. In the text, works are cited as follows: authors' last name and date, without comma, in parentheses: for example, (Jones 1987); with two authors: (Jones and Freeman 1973); with more than two: (Jones et al. 1985); with more than one source cited together (Jones 1987; Freeman 1986); with two or more works by one author: (Jones 1985, 1987).
2. Unless confusion would result, do not use "p." or "pp." before page numbers: for example, (Jones 1987, 115).
3. When the reference list contains more than one work of an author, the suffix *a*, *b*, etc. follows the date in the text citation: for example, (Jones 1987a) or (Jones 1987a; Freeman 1985b).
4. If author's name is mentioned in the text, it need not be repeated in the citation; for example, "Jones (1987, 115) says. . . ."
5. Citations to institutional works should use acronyms or short titles where practicable; for example, (AAA ASOBAT 1966); (AICPA Cohen Commission Report 1977). Where brief, the full title of an institutional work might be shown in a citation: for example, (ICAEW *The Corporate Report* 1975).
6. If the manuscript refers to statutes, legal treatises, or court cases, citations acceptable in law reviews should be used.

Reference List: Every manuscript must include a list of references containing only those works cited. Each entry should contain all data necessary for unambiguous identification. With the author-date system, use the following format recommended by the *Chicago Manual*:

1. Arrange citations in alphabetical order according to surname of the first author or the name of the institution responsible for the citation.
2. Use author's initials instead of proper names.
3. Dates of publication should be placed immediately after author's name.
4. Titles of journals should not be abbreviated.
5. Multiple works by the same author(s) should be listed in chronological order of publication. Two or more works by the same author(s) in the same year are distinguished by letters after the date.
6. Inclusive page numbers are treated as recommended in *Chicago Manual* section 8.67.

Sample entries are as follows:

American Accounting Association, Committee on Concepts and Standards for External Financial Reports. 1977. *Statement on Accounting Theory and Theory Acceptance*. Sarasota, FL: AAA.

Fabozzi, F., and I. Pollack, eds. 1987. *The Handbook of Fixed Income Securities*. 2d ed. Homewood, IL: Dow Jones-Irwin.

Demski, J. S., and D. E. M. Sappington. 1989. Hierarchical structure and responsibility accounting. *Journal of Accounting Research* 27 (Spring): 40–58.

Dye, R., B. Balachandran, and R. Magee. 1989. Contingent fees for audit firms. Working paper, Northwestern University, Evanston, IL.

Kahneman, D., P. Slovic, and A. Tversky, eds. 1982. *Judgment Under Uncertainty: Heuristics and Biases*. Cambridge, United Kingdom: Cambridge University Press.

Porcano, T. M. 1984a. Distributive justice and tax policy. *The Accounting Review* 59 (October): 619–36.

——. 1984b. The perceived effects of tax policy on corporate investment intentions. *The Journal of the American Taxation Association* 6 (Fall): 7–19.

Shaw, W. H. 1985. Empirical evidence on the market impact of the safe harbor leasing law. Ph.D. dissertation, University of Texas, Austin.

Sherman, T. M., ed. 1984. *Conceptual Framework for Financial Accounting*. Cambridge, MA: Harvard Business School.

Footnotes: Footnotes are not used for documentation. Textual footnotes should be used only for extensions and useful excursions of information that if included in the body of the text might disrupt its continuity. Footnotes should be double spaced and numbered consecutively throughout the manuscript with superscript Arabic numerals. Footnotes are placed at the end of the text.

SUBMISSION OF MANUSCRIPTS

Authors should note the following guidelines for submitting manuscripts:

1. Manuscripts currently under consideration by another journal or other publisher should not be submitted. The author must state that the work is not submitted or published elsewhere.
2. In the case of manuscripts reporting on field surveys or experiments, four copies of the instrument (questionnaire, case, interview plan, or the like) should be submitted.
3. Four copies should be submitted together with a check in U.S. funds for $50.00 for members or $100.00 for nonmembers of the AAA made payable to the American Accounting Association. Effective January 1990, the submission fee is nonrefundable.
4. The author should retain a copy of the paper.
5. Revisions must be submitted within 12 months from request, otherwise they will be considered new submissions.

COMMENTS

Comments on articles previously published in *The Accounting Review* will be reviewed (anonymously) by two reviewers in sequence. The first reviewer will be the author of the original article being subjected to critique. If substance permits, a suitably revised comment will be sent to a second reviewer to determine its publishability in *The Accounting Review*. If a comment is accepted for publication, the original author will be invited to reply. All other editorial requirements, as enumerated above, also apply to proposed comments.

POLICY ON REPRODUCTION

An objective of *The Accounting Review* is to promote the wide dissemination of the results of systematic scholarly inquiries into the broad field of accounting.

Permission is hereby granted to reproduce any of the contents of the *Review* for use in courses of instruction, as long as the source and American Accounting Association copyright are indicated in any such reproductions.

Written application must be made to the Editor for permission to reproduce any of the contents of the *Review* for use in other than courses of instruction—e.g., inclusion in books of readings or in any other publications intended for general distribution. In consideration for the grant of permission by the *Review* in such instances, the applicant must notify the author(s) in writing of the intended use to be made of each reproduction. Normally, the *Review* will not assess a charge for the waiver of copyright.

Except where otherwise noted in articles, the copyright interest has been transferred to the American Accounting Association. Where the author(s) has (have) not transferred the copyright to the Association, applicants must seek permission to reproduce (for all purposes) directly from the author(s).

POLICY ON DATA AVAILABILITY

The following policy has been adopted by the Executive Committee in its April 1989 meeting.

"An objective of (*The Accounting Review, Accounting Horizons, Issues in Accounting Education*) is to provide the widest possible dissemination of knowledge based on systematic scholarly inquiries into accounting as a field of professional, research, and educational activity. As part of this process, authors are encouraged to make their data available for use by others in extending or replicating results reported in their articles. Authors of articles which report data dependent results should footnote the status of data availability and, when pertinent, this should be accompanied by information on how the data may be obtained."

OCCASIONAL REVIEWERS AND ASSOCIATE EDITORS

The following faculty members have graciously donated their time and expertise to assist in evaluating manuscripts for *The Accounting Review* between June 1990 and November 1991. Their support is greatly appreciated. The Editor.

Jeffery Abarbanell, *University of Michigan*
Mohammad J. Abdolmohammadi, *Bentley College*
Margaret A. Abernethy, *The University of Melbourne*
Noel Addy, *Mississippi State*
Anwer Ahmed, *University of Florida*
Bipin Ajinkya, *University of Florida*
Mimi Alciatore, *Southern Methodist University*
Tarek Amer, *University of Florida*
Amin H. Amershi, *University of Minnesota*
Brenda H. Anderson, *Boston University*
Stephen Asare, *University of Florida*
Alison Ashton, *Duke University*
Anthony A. Atkinson, *Mount Allison University*
Vidya Awasthi, *Santa Clara University*
William Baber, *George Washington University*
Joseph Bachar, *Tel Aviv University*
Stephen Baginski, *Florida State University*
Paul Bahnson, *University of Montana*
Charles D. Bailey, *Central Florida University*
Stanley Baiman, *Carnegie Mellon University*
K. R. Balachandran, *New York University*
Ramamurthy Balakrishnan, *University of Iowa*
Michael Bamber, *University of Georgia*
Satisprasad Bandyopadhya, *University of Waterloo*
Rajiv D. Banker, *University of Minnesota*
Da-Hsien Bao, *Rutgers University*
Charles Baril, *James Madison University*
Sasson Bar-Yosef, *The Hebrew University of Jerusalem*
Mary E. Barth, *Harvard University*
Jon W. Bartley, *North Carolina State University*
Eliahu Bartov, *University of Rochester*
Allen W. Bathke, Jr., *Florida State University*
Randolph Beatty, *The University of Chicago*
William Beaver, *Stanford University*
Gordon Bechtel, *University of Florida*
Paul Beck, *University of Illinois*
Jean C. Bedard, *University of Connecticut*
Itzak Benbasat, University of British Columbia
Sanford Berg, *University of Florida*
Stanley Biggs, *University of Connecticut*
Sarah Bonner, *University of Colorado–Boulder*
Efrim Boritz, *University of Waterloo*
Rien Bouwman, *University of North Carolina*
Robert Bowen, *University of Washington*
Robert G. Bowman, *University of Auckland*
Phelim P. Boyle, *University of Waterloo*
Robert Bricker, *Case Western Reserve University*
Clifton Brown, *University of Illinois–Urbana-Champaign*
Lawrence Brown, *SUNY at Buffalo*
Paul Brown, *New York University*
Philip R. Brown, *University of Western Australia*
Bruce Bublitz, *University of Kansas*
Thomas Buchman, *University of Colorado at Boulder*
Frank Buckless, *North Carolina State University*
David Burgstahler, *University of Washington*
Robert Bushman, *The University of Chicago*
Richard P. Castanias, *University of California–Davis*
Peter Chalos, *University of Illinois at Chicago*
Ramesh Chandra, *University of Windsor*
Paul K. Chaney, *Vanderbilt University*
John J. Cheh, *New York University*
Kevin Chen, *Rutgers University*
Agnes Cheng, *University of Houston*
Peter Cheng, *Purdue University*
Joseph K. Cheung, *George Mason University*
Charles Chi, *University of Southern California*
Frederick D. S. Choi, *New York University*
Freddie Choo, *San Francisco State University*
Chee Chow, *San Diego State University*
Mary York Christ, *Pennsylvania State University*
Bryan Church, *Wilfrid Laurier University*
Susan Cohen, *University of Illinois–Champaign*
Daniel Collins, *University of Iowa*
Jean C. Cooper, *University of Kentucky*
Dean Crawford, *Oakland University*
William Cready, *Texas A&M*
Andrew Cuccia, *Louisiana State University*
Lane Daley, *University of Minnesota*
Somnath Das, *University of California–Berkeley*
Henry Davis, *Indiana University*
Jon Davis, *University of Colorado–Boulder*
Michael L. Davis, *Lehigh University*
Prabudha De, *The University of Dayton*
Edward Deakin, *University of Texas at Austin*
Mark L. DeFond, *University of Southern California*
Dan S. Dhaliwal, *University of Arizona*
Bala Dharan, *Rice University*
J. Richard Dietrich, *University of Illinois*
William Dilla, *University of Illinois*
Jesse F. Dillard, *University of New Mexico*
Michael Duffy, *University of Southern California*
Amitabh Dugar, *Michigan State University*
Gordon Leon Duke, *University of Minnesota*
Roland (Pete) Dukes, *University of Washington*
Peter D. Easton, *MacQuarrie University*
Samir El-Gazzar, *Rutgers University*
Kamal El-Sheshai, *Georgia State University*
Pieter Elgers, *University of Massachusetts*
Rachamin Elitzur, *University of Toronto*
Dan Elnathan, *University of Southern California*
Kirsten M. Ely, *University of California–Los Angeles*
Craig Emby, *Simon Fraser University*

Robert D. Emerson, *University of Florida*
Charles Enis, *Pennsylvania State University*
Pamela E. Erickson, *Vanderbilt University*
Mahmoud Ezzamel, *Manchester Institute of Science & Technology*
Patricia M. Fairfield, Georgetown University
Jerry Feltham, *University of British Columbia*
Ehsan H. Feroz, *University of Minnesota–Duluth*
Kenneth Ferris, *Southern Methodist University*
David Finley, *Simon Fraser University*
Mark Flannery, *University of Florida*
Richard K. Fleischman, *John Carroll University*
George Foster, *Stanford University*
Taylor W. Foster, III, *New Mexico State University*
Jennifer Francis, *Duke University*
Jere Francis, *University of Iowa*
Werner G. Frank, *University of Wisconsin–Madison*
David Frederick, *University of Colorado–Boulder*
James R. Frederickson, *Indiana University*
Robert Freeman, *University of Texas–Austin*
Carol Frost, *Washington University*
Veronique Frucot, *Syracuse University*
John C. Gardner, *SUNY–Binghamton*
Jennifer J. Gaver, *University of Georgia*
Kenneth M. Gaver, *University of Georgia*
Dimitrios Ghicas, *CUNY–Binghamton*
Michael Gift, *University of South Florida*
Gary Giroux, *Texas A&M University*
G. William Glezen, *University of Arkansas*
James T. Godfrey, *George Mason University*
Joseph Godwin, *University of Georgia*
Stephen R. Goldberg, *Purdue University*
Suresh Govindaraj, *New York University*
V. Govindarajan, *Dartmouth College*
Julia Grant, *Case Western Reserve University*
Sidney Gray, *University of Glasgow*
Sharon L. Green, *University of Wisconsin–Madison*
Paul Griffin, *University of California–Davis*
Sanjay Gupta, *Arizona State University*
Karl Hackenbrack, *University of Florida*
Susan Haka, *Michigan State University*
Thomas Hall, *University of Texas–Arlington*
James R. Hamill, *University of Oklahoma*
Jonathan Hamilton, *University of Florida*
John Hand, *University of Chicago*
Stephen C. Hansen, *Stanford University*
Paul Harrison, *University of South Carolina*
John Hassell, *University of Texas at Arlington*
In-Mu Haw, *Texas Christian University*
Carla K. Hayn, *Northwestern University*
James Heian, *Pennsylvania State University*
Vicky Heiman-Hoffman, *University of Pittsburgh*
Richard Helleloid, *University of Wisconsin–Eau Claire*
William Hillison, *Florida State University*
Mark Hirst, *University of New South Wales*
Peggy Hite, *Indiana University*
Joanna L. Ho, *University of California–Irvine*
Doris Holt, *Iowa State University*
William Hopwood, *University of Houston*
Keith Houghton, *University of Melbourne*
Carol Houston, *San Diego State University*
Patricia Hughes, *University of Southern California*
Wesley Hutchinson, *University of Florida*
Yuchang Hwang, *University of Pittsburgh*
Yuji Ijiri, *Carnegie Mellon University*
Eugene Imhoff, Jr., *University of Michigan*
Robert Ingram, *University of Alabama*
Betty Jackson, *University of Colorado–Boulder*
Christopher James, *University of Florida*
Robert E. Jensen, *Trinity University*
James Jiambalvo, *University of Washington*
Kose John, *New York University*
W. Bruce Johnson, *University of Iowa*
Jennifer Jones, *The University of Chicago*
James S. Jordan, *University of Minnesota*
Woon-Oh Jung, *University of Illinois*
Steven Kachelmeier, *University of Texas at Austin*
Sok-Hyon Kang, *Carnegie Mellon University*
Steven E. Kaplan, *Arizona State University*
Duane Kennedy, *University of Waterloo*
Thomas Kida, *University of Massachusetts*
Jae-Oh Kim, *The Ohio State University*
Oliver Kim, *University of California–Los Angeles*
Raymond King, *University of Oregon*
Ronald King, *Washington University*
William Kinney, Jr., *University of Texas–Austin*
Alison J. Kirby, *University of California–Berkeley*
Don Kleinmuntz, *University of Illinois*
Michael Knapp, *University of Oklahoma*
Lisa Koonce, *University of Texas–Austin*
Leslie Kren, *University of Wisconsin–Milwaukee*
Murugappa Krishnan, *Purdue University*
Jack L. Krogstad, *Creighton University*
William Kross, *Purdue University*
Joseph Lakonishok, *University of Illinois*
Wayne Landsman, *University of North Carolina*
William Lanen, *University of Michigan*
Amy H. Lau, *Oklahoma State University*
Marc J. LeClere, *University of Maryland–College Park*
Changwoo Lee, *New York University*
Charles M. C. Lee, *University of Michigan*
Robert Alan Leitch, *University of South Carolina*
Haim Levy, *University of Florida/Hebrew University*
Steven Lilien, *CUNY–Baruch*
Stephen Limberg, *University of Texas–Austin*
Frederick Lindahl, *Duke University*
Thomas Linsmeier, *University of Iowa*
Robert Lipe, *University of Michigan*
Gerald Lobo, *Syracuse University*
Kenneth Lorek, *Florida State University*
Russell Lundholm, *Stanford University*
Tom Lys, *Northwestern University*
Alan Macnaughton, *University of Waterloo*
G. S. Maddala, *University of Florida*
P. Merle Maddocks, *Memphis State University*
Joseph Magliolo, III, *University of Pennsylvania*
Laureen Maines, *Duke University*
James Manegold, *University of Southern California*
M. Laurentius Marais, *University of Chicago*

Linda McDaniel, *University of Washington*
Gary McGill, *University of Florida*
Morris McInnes, *Suffolk University*
Janet Meade, *University of Houston*
Gary K. Meek, *Oklahoma State University*
Nahum Melumad, *Stanford University*
Krishnagopal Menon, *Boston University*
Yaw M. Mensah, *Rutgers University*
Rayman D. Meservy, *Brigham Young University*
James Millar, *University of Arkansas*
Bruce Miller, *University of California–Los Angeles*
H. Fred Mittelstaedt, *Arizona State University*
Cindy Moeckel, *Arizona State University*
Rosanne M. Mohr, *Portland State University*
Dale Morse, *Cornell University*
Sanford Morton, *University of British Columbia*
Donald Moser, *University of Pittsburgh*
O. Doug Moses, *Naval Postgraduate School*
Stephan Motowidlo, *University of Florida*
Ari Mukerji, *University of Minnesota*
Charles Mulford, *Georgia Institute of Technology*
Dennis Murray, *University of Colorado–Denver*
Mark W. Nelson, *Cornell University*
Harry Newman, *University of Illinois–Chicago*
Paul Newman, *University of Texas–Austin*
Donald R. Nichols, *Texas Christian University*
George E. Nogler, *Suffolk University*
John O'Brien, *Carnegie Mellon University*
Patricia O'Brien, *University of Michigan*
James A. Ohlson, *Columbia University*
Daniel O'Leary, *University of Southern California*
Cherie O'Neill, *University of South Florida*
Chris Olsen, *University of Texas–Austin*
Tae-Young Paik, *New York University*
Kurt Pany, *Arizona State University*
Mohinder Parkash, *Oakland University*
Victor Pastena, *SUNY at Buffalo*
James Patell, *Stanford University*
Evelyn R. Patterson, *University of Michigan*
Yoram Peles, *Hebrew University of Jerusalem*
James M. Peters, *Carnegie Mellon University*
Kathy R. Petroni, *Michigan State University*
Morton Pincus, *Washington University*
R. David Plumlee, *Kansas State University*
Grace Pownall, *Washington University*
Eric G. Press, *SUNY at Buffalo*
Michael Prietula, *Carnegie Mellon University*
Madhav Rajan, *University of Pennsylvania*
Ram T. S. Ramakrishnan, *Columbia University*
Kris Raman, *University of North Texas*
K. Ramesh, *Northwestern University*
Graeme Rankine, *Rice University*
Judy Rayburn, *University of Minnesota*
Zelma Rebman-Huber, *Simon Fraser University*
Philip R. Regier, *Arizona State University*
Stefan Reichelstein, *University of California–Berkeley*
Jane L. Reimers, *Florida State University*
Sarah Reiter, *SUNY–Binghamton*
J. Hal Reneau, *Arizona State University*
Richard D. Rennie, *University of Regina*
Alan J. Richardson, *Queen's University*
Gordon D. Richardson, *University of Waterloo*
David Ricchiute, *Notre Dame*
Byung Ro, *Purdue University*
Gita Roa, *University of Illinois*
Walter Robbins, *University of Alabama*
Robin Roberts, *University of Missouri–Columbia*
Kermit J. Rohrbach, *University of Oklahoma*
Alan Roshwalb, *Georgetown University*
Marc A. Rubin, *Miami University*
Stephen Ryan, *New York University*
William Salatka, *University of Kansas*
Richard Sansing, *Yale University*
David Sappington, *University of Florida*
Bharat Sarath, *University of Pennsylvania*
Shahrokh M. Saudagaran, *Santa Clara University*
Michael Schadewald, *University of Texas–Austin*
Thomas Schaefer, *Florida State University*
Jeffrey Schatzberg, *University of Arizona*
Arnold Schneider, *Georgia Institute of Technology*
Hein Schreuder, *Rijksuniversiteit Limburg*
Barbara W. Scofield, *University of Kentucky*
Thomas Scott, *University of Alberta*
Stephan Sefcik, *University of Washington*
Frank Selto, *University of Colorado–Boulder*
Timothy L. Shaftel, *University of Kansas*
Philip B. Shane, *Penn State University*
Wayne H. Shaw, *University of Colorado–Boulder*
Mohamed Shehata, *McMaster University*
Marjorie K. Shelley, *University of Illinois*
Terry Shevlin, *University of Iowa*
Toshi Shibano, *University of California–Berkeley*
Keith Shriver, *Arizona State University*
Daniel Simon, *Indiana University/Northwest*
Dan Simunic, *University of British Columbia*
Konduru Sivaramakrishnan, *Carnegie Mellon University*
Douglas Skinner, *University of Michigan*
Stephen Slutsky, *University of Florida*
Abbie Smith, *University of Chicago*
Julie Sterner Sobery, *Southern Illinois University*
Lisa Gilbert Soo, *Boston College*
Eric Spires, *The Ohio State University*
Sri Sridharan, *Northwestern University*
Bin Srinidhi, *Rutgers University*
Kent St. Pierre, *James Madison University*
Paul Steinbart, *Memphis State University*
Jens A. Stephan, *University of Cincinnati*
Earl K. Stice, *Rice University*
James Stice, *Brigham Young University*
Scott E. Stickel, *University of Pennsylvania*
Thomas Stober, *Indiana University*
Don J. Stokes, *University of Queensland*
Sivakumar Swaminathan, *Michigan State University*
Edward Swanson, *Texas A&M University*
Walter Teets, *University of Illinois*
Siew Hong Teoh, *University of California–Los Angeles*
Sundaraman Thiagarajan, *Tulane/Northwestern University*
Lynda Thoman, *Purdue University*

Jacob Thomas, *Columbia University*
Robert Thompson, III, *University of Maryland*
Mikel G. Tiller, *Indiana University*
Robert H. Trezevant, *University of Southern California*
Mark Trombley, *University of Arizona*
Gregory M. Trompeter, *University of Notre Dame*
Brett Trueman, *University of California–Berkeley*
Richard Tubbs, *University of Iowa*
Robert Tucker, *University of Illinois at Chicago*
Carol Venable, *San Diego State University*
P. C. Venkatesh, *University of Cincinnati*
Robert E. Verrecchia, *University of Pennsylvania*
Robert L. Vigeland, *Texas Christian University*
James Waegelein, *University of Kansas*
David Wallin, *The Ohio State University*
Shiing-wu Wang, *University of Arizona*
Charles Wasley, *Washington University*
John S. Watts, *Purdue University*
Ron Weber, *University of Queensland*
Joseph Weintrop, *SUNY at Buffalo*
Robert Whaley, *Duke University*
David Williams, *The Ohio State University*
Arnold Wright, *University of New South Wales*
Jeffrey Yost, *University of Florida*
Mark Young, *University of Colorado–Boulder*
Richard Young, *The Ohio State University*
Paul Zarowin, *New York University*
Christine Zavgren, *Clarkson University*
Daniel Zeghal, *University of Ottawa*
David Ziebart, *University of Illinois*
Jerold L. Zimmerman, *University of Rochester*
Mark E. Zmijewski, *University of Chicago*

THE ACCOUNTING REVIEW
Vol. 67, No. 2
April 1992
pp. 249-283

Belief-Function Formulas for Audit Risk

Rajendra P. Srivastava
Glenn R. Shafer
University of Kansas

SYNOPSIS AND INTRODUCTION: This article relates belief functions to the structure of audit risk and provides formulas for audit risk under certain simplifying assumptions. These formulas give plausibilities of error in the belief-function sense.

We believe that belief-function plausibility represents auditors' intuitive understanding of audit risk better than ordinary probability. The plausibility of a statement, within belief-function theory, measures the extent to which we lack evidence against the statement. High plausibility for error indicates only a lack of assurance, not positive evidence that there is error. Before collecting, analyzing, and aggregating the evidence, an auditor may lack any assurance that a financial statement is correct, and in this case will attribute very high plausibility to material misstatement. This high plausibility does not necessarily indicate any evidence that the statement is materially misstated, and hence, it is inappropriate to interpret it as a probability of material misstatement.

The SAS No. 47 formula for audit risk is based on a very simple structure for audit evidence. The formulas we derive in this article are based on a slightly more complex but still simplified structure, together with other simplifying assumptions. We assume a tree-type structure for the evidence, assume that all evidence is affirmative and that each variable in the tree is binary. All these assumptions can be relaxed. As they are relaxed, however, the formulas become more complex and less informative, and it then becomes more useful to think in terms of computer algorithms rather than in terms of formulas (Shafer et al. 1988).

In general, the structure of audit evidence corresponds to a network of variables. We derive formulas only for the case in which each item of

This research has been supported in part by grants from the Peat Marwick Foundation, the General Research Fund of the University of Kansas, and the National Science Foundation (grant IRI8902444). The authors would like to thank Andrew D. Bailey, Jr., Peter Gillett, Lynford E. Graham, William R. Kinney, Jr., Prakash P. Shenoy, W. R. Scott, Ira Solomon, and John Willingham for valuable comments on an earlier version of the article.

Submitted July 1989.
Accepted August 1991.

evidence bears either on all the audit objectives of an account or on all the accounts in the financial statement, as in figure 1, so that the network is a tree. Usually, however, there will be some evidence that bears on some but not all objectives for an account, on some but not all accounts, or on objectives at different levels; in this case, the network will not be a tree.

We assume that all evidence is affirmative because this is the situation treated by the SAS No. 47 formula and because belief-function formulas become significantly more complex when affirmative and negative evidence is combined. This complexity is due primarily to the renormalization involved in Dempster's rule for combining belief functions.

The variables in the network or tree represent various audit objectives, accounts, and the financial statement as a whole. We assume these variables are binary. For example, we assume that an account either is or is not materially misstated. This assumption is clearly too restrictive for most audit practice. Often, for example, an auditor must consider immaterial errors in individual accounts that could produce a material error in the financial statement when they are aggregated.

We derive formulas for plausibility of material misstatement at three levels: the financial statement level, the account level, and the audit objective level. The formula at the audit objective level resembles the SAS No. 47 formula,[1] but the formulas at the other two levels are significantly different. Because our model does distinguish evidence gathered at the three different levels, audits based on our formulas are sometimes significantly more efficient[2] than audits based on the SAS No. 47 model or on the simpler Bayesian models.

Key Words: *Audit risk, Belief functions, Planning model, Evaluation model.*

---

[1] With the evidential structure that we consider in this article, we obtain plausibility formulas that are closer in form to the SAS No. 47 formula than to Bayesian formulas (Leslie 1984). This is because the belief-function theory permits an auditor to have belief, say 0.6, based on the procedures performed, that a given objective is met, without having to assign the remaining 0.4 of his or her belief to whether the objective has been met or not. This unassigned belief, 0.4, in the case, represents the plausibility that the objective is not met.

Those accustomed to Bayesian thinking may prefer to express judgments about the effectiveness of procedures in terms of conditional probabilities for detection of error, given the presence or absence of such error. In our view, such conditional probabilities are only one way of expressing intuitive judgments that may alternatively be expressed in terms of belief functions. We would express our intuitive judgments as conditional probabilities only if we intended to carry out a completely Bayesian analysis. We would express our intuitive judgments directly in terms of *m*-values if we intended to carry out a belief-function analysis. It is possible, nonetheless, to express intuitive judgments in terms of conditional probabilities and then to translate these conditional probabilities into belief functions. There is some arbitrariness involved in the translation, but it can be carried out in such a way that the Bayesian approach becomes a special case of the belief-function approach (Shafer 1982). If we combine the belief function representing the conditional probabilities with a Bayesian prior using Dempster's rule, we will obtain the appropriate Bayesian posterior.

[2] In general, the more accurately we model the structure of audit evidence, the more effective and efficient we can expect the audit to be. However, it is also important to recognize that if the necessary inputs to a more accurate model cannot be estimated accurately by the auditor, decreases in audit effort mandated by the model can decrease audit effectiveness.

THE remainder of this article is divided into seven sections and two appendices. In section I, we review the existing literature on the audit-risk model. In section II, we review the structure of various types of audit evidence. In section III, we review the belief-function treatment of audit evidence and discuss using belief-function plausibility to represent audit risk. In section IV, we discuss the combination of beliefs (or m-values) at each level of the financial statement. This combination corresponds to aggregating the evidence that bears directly on the objectives at those levels. In section V, we present total plausibility formulas (i.e., belief-function formulas for the total audit risk) and numerical examples. In section VI, we highlight the limitations of our formulas and suggest future directions for research. In section VII, we summarize our results. We review the propagation of belief functions in a tree in appendix A, and we use these results to derive the belief-function formulas for the total audit risk in appendix B.

I. Review of the Literature

Here we will review the existing literature on audit risk formulas and the difficulties involved in interpreting the numbers in these formulas as probabilities.

Although the model of SAS No. 47 (AICPA 1983) has been used for planning purposes since June 1984, controversy about its applicability for planning and evaluation has persisted. Graham points out, for example, that "overall assessments of audit risk for the financial statements taken as a whole are usually impractical for audit planning and engagement control" (1985a, 14), since inherent and control risks can vary from one account to another and from one class of transactions to another. Graham further suggests that the audit-risk model should be decomposed into components that determine audit risk for management assertions related to each account balance (accounts receivable, inventory, etc.) or class of transactions (purchases, sales, etc.). Cushing and Loebbecke (1983) argue that the SAS No. 47 model provides no guidance on aggregating risks associated with various account balances, and Kinney (1989) has demonstrated recently that the model has properties that may significantly understate achieved audit risk.

Bayesian models have also been discussed in the literature for risk aggregation (see, e.g., Kinney 1984; Leslie 1984). The simplest Bayesian models do not consider different levels of the account, but Boritz and Jensen (1985) discuss the hierarchical structure of audit evidence and propose ways to combine such evidence by using probabilities to represent risks. Also, in discussing their assertion-based approach to auditing, Leslie et al. (1986) recognized the importance of the structure of audit evidence and emphasized that assurances from various items of evidence should be assessed at the management assertion level of the account and then combined. They considered the relationship between various accounts (e.g., accounts receivable depends on sales and cash receipts) in the aggregation process, but they assumed that the different items of evidence for different accounts were independent.[3]

At a qualitative level, there has been increasing recognition of the importance of the structure of evidence and its relation to various audit objectives in assessing control risk. In SAS No. 55 entitled "The Auditor's Responsibility for Assessing Control Risk" (AICPA 1988a, par. 3), we find the following statement:

---

[3] Leslie et al. (1986) assume Poisson error rates, and they base their analysis on assurance factors, which represent Poisson parameters. They do not explicitly assume the independence of different items of evidence, but this assumption appears to be the justification for their practice of summing individual assurance factors to obtain an overall level.

> After obtaining this understanding, the auditor assesses control risk for the assertions embodied in the account balance, transaction class, and disclosure components of the financial statements.

Although the AICPA has not yet required auditors to consider individual audit objectives when assessing other risks (e.g., inherent risk and detection risk), it appears that many are already doing so (Graham 1985a–1985e).

Boritz and Wensley (1990) have used the structure of audit evidence to develop their computer system for audit planning and evaluation, but their system uses heuristic rules rather than formal theory to combine uncertainties. Since such heuristics have been seen to fail in complex systems (Buchanan and Shortliffe 1984), we believe that more needs to be done to develop theoretical methods of combination that take account of the complexity of the structure of audit evidence.

We advance belief functions as a basis for this theoretical development because we believe that the usefulness of the Bayesian approach is limited by divergences between the intuitive and Bayesian interpretations of audit risk. For example, according to SAS No. 47, if an auditor decides not to consider inherent factors, then the inherent risk is set equal to 1. Since a probability of 1 means certainty, this seems to be saying that it is certain that the account is materially in error. But this is not what the auditor has in mind when deciding not to depend on inherent factors. The auditor's intention is represented better by a belief-function plausibility of 1 for material error, which says only that the auditor lacks evidence based on inherent factors.

In a less extreme situation, the auditor may believe, on the basis of inherent factors, that the account is fairly stated and yet be unwilling to rely on these factors past a certain point. In this case, the auditor may, as SAS No. 47 suggests, assign a value less than the maximum, say 70 percent, to inherent risk. If interpreted in probability terms, this number says that the inherent factors give a 30 percent chance that the account is not materially misstated and a 70 percent chance that it is materially misstated. This suggests that the evidence is negative, contrary to the auditor's intuition. The probability interpretation is even more confusing if the auditor sets the inherent risk at 50 percent. What does this mean? Does it mean that the auditor is completely ignorant about the state of the account, or does it mean there is more evidence that the account is not being materially misstated than when only 30 percent assurance was assumed?

Belief functions, since they permit uncommitted belief, allow us to interpret the auditor's choices in a straightforward way. When the auditor sets the risk at 70 percent, a 30 percent degree of support from inherent factors is claimed, leaving 70 percent of the auditor's belief uncommitted. In this case, material misstatement has 70 percent plausibility, but absence of material misstatement has 100 percent plausibility. When the auditor sets the risk at 50 percent, stronger support from inherent factors is claimed. In this case, the plausibility of material misstatement is reduced to 50 percent, while the plausibility of no material misstatement remains at 100 percent.

We believe that belief functions provide a flexible and adaptable way to combine evidence from a variety of sources (Akresh et al. 1988). One aspect of this flexibility is that, when the belief functions representing individual items of evidence are probability measures, the belief-function analysis reduces to a Bayesian analysis (Shafer and Srivastava 1990).

II. The Structure of Audit Evidence

As we have seen, neglect of the structure of audit evidence has been a problem in some risk models. Before developing belief-function formulas we must, therefore, specify carefully the kind of structure we are considering.

We will adopt the structure currently assumed in auditing standards (AICPA 1988a; see also, e.g., Arens and Loebbecke 1988), with some simplifications. The standards generally divide audit evidence into four categories: (1) general knowledge about inherent risk, (2) evidence from analytical procedures, (3) knowledge of control factors and accounting systems, and (4) tests of details of balances. Within each general category, further structure arises because of the relevance of different items of evidence to different accounts and different objectives. In the following subsections, we review the structure within each of the four categories.

General Knowledge about Inherent Risk

In this category, we include general knowledge about risk factors that lie outside of the accounting system and also outside of the auditor's control. Examples include economic, political, business and regulatory environments, experience from the prior year's audit, management philosophy and style, organizational structure, and audit committee (see SAS No. 55 for more examples). Also included in this category are factors that make individual accounts more or less susceptible to error, such as the complexity of transactions, the volume of transactions processed, the susceptibility of assets to defalcation, and related party transactions. Such factors are important for the auditor's planning decisions, since the extent, nature, and timing of tests will depend on the auditor's assessment of the effect of these factors on the individual accounts and on the financial statement as a whole.

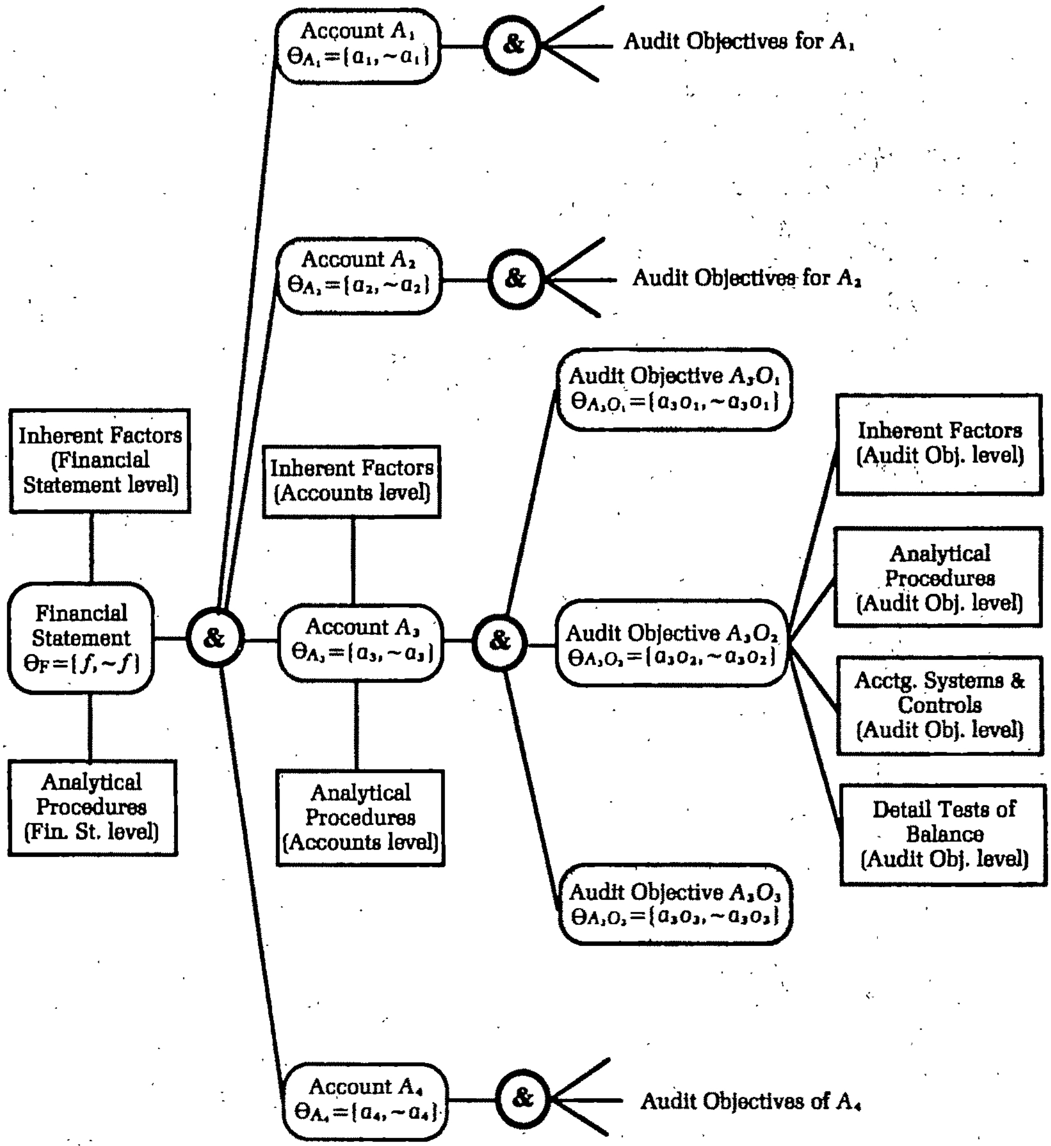
Some inherent factors affect entire financial statements, whereas others affect only certain accounts or classes of transactions. Some affect only a particular audit objective for an account or a class of transactions. Information about the competence and integrity of management, for example, will affect the entire financial statement. The auditor will have a higher level of assurance about the financial statement when management is of recognized competence and integrity than when management is known to have been involved previously in irregularities. In contrast, an auditor auditing a newspaper publisher realizes that libel suits against newspapers are common and will treat this knowledge as evidence affecting only accrued-contingent liability. The nature and complexity of an individual account (e.g., susceptibility to defalcation, volume of transactions, non-routine transactions, account balance based on management's judgment) also may affect only that account. As an example of evidence that affects only a certain audit objective, we might cite information about related party transactions from the minutes of board meetings. This information affects only the disclosure and classification objectives of the accounts involved in the transaction.

In summary, we see that evidence about inherent factors can bear on the financial statement at three different levels: (1) the financial statement level, (2) the individual account or class of transactions level, and (3) the audit objective level for individual accounts or classes of transactions. See figure 1 for details.

Analytical Procedures

Recently, the AICPA published SAS No. 56 (AICPA 1988b), requiring auditors to use analytical procedures on all audit engagements. According to SAS No. 56 (par. 2),

Figure 1
An Evidential Network


Account A_1
$\Theta_{A_1}=\{a_1,\sim a_1\}$
&
Audit Objectives for A_1
Account A_2
$\Theta_{A_2}=\{a_2,\sim a_2\}$
&
Audit Objectives for A_2
Audit Objective A_3O_1
$\Theta_{A_3O_1}=\{a_3o_1,\sim a_3o_1\}$
Inherent Factors (Financial Statement level)
Inherent Factors (Accounts level)
Inherent Factors (Audit Obj. level)
Analytical Procedures (Audit Obj. level)
Financial Statement
$\Theta_F=\{f,\sim f\}$
&
Account A_3
$\Theta_{A_3}=\{a_3,\sim a_3\}$
&
Audit Objective A_3O_2
$\Theta_{A_3O_2}=\{a_3o_2,\sim a_3o_2\}$
Acctg. Systems & Controls (Audit Obj. level)
Analytical Procedures (Fin. St. level)
Analytical Procedures (Accounts level)
Detail Tests of Balance (Audit Obj. level)
Audit Objective A_3O_3
$\Theta_{A_3O_3}=\{a_3o_3,\sim a_3o_3\}$
Account A_4
$\Theta_{A_4}=\{a_4,\sim a_4\}$
&
Audit Objectives of A_4

Note: A rounded rectangle represents a variable (variables being the financial statements as a whole, various accounts, and the related audit objectives). A rectangle represents an item of audit evidence. The evidence is connected to a variable that it directly supports. A circle with "&" implies that the variable on the left is true if and only if the variables on the right of the circle are true.

> Analytical procedures are an important part of the audit process and consist of evaluations of financial information made by a study of plausible relationships among both financial and non-financial data. Analytical procedures range from simple comparisons to the use of complex models involving many relationships and elements of data.

The statement proposes that analytical procedures be used for the following purposes (par. 4):

1. To assist the auditor in planning the nature, timing and extent of other auditing procedures.
2. As a substantive test to obtain evidential matter about particular assertions related to account balances or class of transactions.
3. As an overall review of the financial information in the final review stage of the audit.

According to SAS No. 56, the expected effectiveness and efficiency of an analytical procedure depends on (1) the nature of the audit objectives, (2) the plausibility and predictability of the relationship, (3) the reliability and availability of the data used to develop the expectation, and (4) the precision of the expectation.

Like general knowledge about inherent factors, analytical procedures can provide assurance at various levels. The more common analytical procedures seem to provide assurance at either the account level or the audit objective level. A comparison of the previous year's accounts payable with the current period's accounts payable provides assurance for the accounts payable balance as a whole. A comparison of the previous year's ratio of bad debt expense to accounts receivable balance with the current year's ratio would bear on collectibility of accounts receivable, a valuation objective.

In the present article, for the purpose of completeness, we will assume analytical procedures to be effective at all three levels: (1) the financial statement level, (2) the account or class of transactions level, and (3) the audit objective level for accounts or classes of transactions. When certain items of evidence are not to be considered in an audit, then those items are eliminated by setting the corresponding plausibilities to 1 (see the discussion in section V).

Control Factors and Accounting Systems

We include in this category all items of evidence related to accounting systems, control procedures, and tests of transactions. A test of controls typically bears on the audit objective level of an individual account, while a test of transactions typically bears on the audit objective level of a class of transactions. Controls built into a cost accounting system, for example, bear on the valuation objective of inventory, while the use of prenumbered bills of lading and sales invoices periodically accounted for bears on the completeness objective of sales.

Tests of Details of Balances

Tests of details of balances bear primarily on the audit objective level. Since it is costly to obtain this type of evidence, the auditor minimizes the need for it by maximizing the assurance to accounts and audit objectives from other sources.

Some tests of details balance bear on only one audit objective, while others may bear on more than one. A review of the minutes of board meetings to check whether receivables have been factored bears only on the ownership objective of accounts receivable, but confirmations of accounts receivable by customers provide assurance for both the existence and valuation objectives. In general, such assurances may vary in strength from objective to objective. For example, confirmation of accounts receivable may provide a higher level of support for the existence objective than for the valuation objective.

When an item of evidence bears equally on *all* the objectives of an account or *all* the accounts of the financial statement, we can represent it within a tree structure by linking it directly to the account or the financial statement, as the case may be. But when a test provides support to more than one audit objective, say, but not equally to all the objectives at once, we obtain a network of variables that is not a tree, and this makes the derivation of formulas cumbersome. The formulas given here are based on the assumption that the network is a tree, but they can be used as approximations in the non-tree case. One way to use them as approximations is to treat the evidence as if it consisted of independent items of evidence bearing on the different objectives. The formula for the total audit risk (i.e., total plausibility of error) at the audit objective level will still be valid (i.e., it maintains its multiplicative form; see section V) when we do this, but the formulas at the account level and the financial statement level will provide only a conservative estimate of the total risk (i.e., plausibility of error). For example, suppose that confirmations of accounts receivable yield 0.9 level of assurance that both existence and valuation objectives are met. If we treat this as two items of evidence, one giving 0.9 degree of support for existence and one giving 0.9 for valuation, then our formulas give a total assurance, for the two objectives jointly, of $0.9\times0.9=0.81$, corresponding to a risk (plausibility of error) of 0.19. But the correct value for the assurance is 0.9, corresponding to a risk of 0.1.

III. Belief-Function Approach to Audit Evidence

In this section, we review the belief-function approach to representing uncertainties in audit evidence. The belief-function framework involves three related representations for beliefs concerning a topic: the belief function (***Bel***), the plausibility function (***PL***), and the basic probability assignment (**m**). As we will explain, the basic probability assignment is often convenient for expressing initial judgments, but the plausibility function is useful for expressing final judgments about audit risk.

The basic probability assignment is also called the **m**-function, and its values are called **m**-values (Shafer 1976). The basic difference between **m**-values and probabilities is that probabilities are assigned to individual elements of a frame,[4] say Θ, whereas **m**-values are assigned to a subset of elements of the frame. The sum of all the **m**-values for all the subsets of the frame Θ is 1. Formally, the **m**-function assigns a number $\mathbf{m}(B)$ to each subset B of Θ such that $\mathbf{m}(\varnothing)=0$($\varnothing$ being the empty set) and:

$$\sum_{B\subseteq\Theta}\mathbf{m}(B)=1.$$

---

[4] We call an exhaustive and mutually exclusive set of possible answers to a question a *frame*. We will often use the symbol Θ to represent the frame in which we are interested. In the case of a yes-no question, the frame has only two elements; $\Theta=\{\text{yes,no}\}$, or $\Theta=\{a=$account 'A' is not materially misstated, $\sim a=$account 'A' is materially misstated$\}$, etc. But, in general, a frame may be a very large set, for its question may have many possible answers.

There are two ways to obtain **m**-values on a frame: (1) they may be assigned directly by the decision maker on the basis of subjective judgment and (2) they may be derived from a compatibility relationship between a frame with known probabilities and the frame of interest.[5] We will use the first approach to discuss our example.

Suppose the auditor has performed a set of analytical procedures appropriate to account 'A' and finds no discrepancy or errors in the account. On the basis of this observation, the auditor feels that the evidence is positive and provides a medium level of support, say 0.6, to 'a' that the account is not materially misstated.[6] However, at the same time, the auditor feels that there is nothing to indicate that the account is materially misstated ($\sim a$). This means that 0.6 degree of support is assigned to 'a', 0 to '$\sim a$', and the remaining 0.4 is the ignorance assigned to the entire frame $\Theta=\{a,\sim a\}$; that is,

$$\mathbf{m}_{PA}(a)=0.6,\ \mathbf{m}_{PA}(\sim a)=0,\ \text{and}\ \mathbf{m}_{PA}(a,\sim a)=0.4,$$

where the subscript PA stands for analytical procedures at the account level. The above set of **m**-values represents affirmative evidence.[7]

As mentioned earlier, we will consider only affirmative evidence in our derivation of audit risk formulas. Although the approach of aggregating evidence discussed in this article is valid for any type of evidence, use of affirmative evidence avoids the renor-

---

[5] In general, if we want to make probability judgments about the elements of a frame Θ for which we have no probability measures, we can do so by relating the elements of Θ to the elements of the frame S for which we have knowledge of its probability distribution. This relationship is called the compatibility relationship. The basic idea is that each probability $P(s)$, where s is an element of S, should contribute to a degree of belief in the subset $\Gamma(s)$ of Θ consisting of elements with which s is compatible. If several s items have the same $\Gamma(s)$—in other words, $\Gamma(s)$ is equal to B for several s items—then the probabilities of all these will contribute to our degree of belief that the answer to the question considered by Θ is somewhere in B. For each subset B of Θ, let $\mathbf{m}(B)$ be the total probability for all the s items whose $\Gamma(s)$ is equal to B:

$$\mathbf{m}(B)=\sum_{\Gamma(s)=B} P(s).$$

It follows from this formula that:

$$\sum_{B\subseteq\Theta}\mathbf{m}(B)=1,$$

and

$$\mathbf{m}(\varnothing)=0,$$

where $\varnothing$ is the empty set.

[6] As a general convention, we will use capital letters to denote names of accounts or audit objectives (nodes) and small letters in script to represent their values.

[7] Affirmative evidence implies that the evidence directly supports the assertion to a certain degree and provides no support for its negation. In our example, we have $\mathbf{m}_{PM}(a)=0.6$, $\mathbf{m}_{PA}(\sim a)=0$, and $\mathbf{m}_{PM}(\{a,\sim a\})=0.4$, which implies that we have direct evidence that 'a' is met with 0.6 degree of support and no evidence that 'a' is not met.

Negative evidence implies that the evidence directly supports the negation of the assertion and provides no support for the assertion, i.e.,

$$\mathbf{m}_{PM}(a)=0,\ \mathbf{m}_{PM}(\sim a)=0.4,\ \text{and}\ \mathbf{m}_{PM}(\{a,\sim a\})=0.6.$$

A mixed item of evidence can be defined as an item that provides some support for the assertion and some for its negation, i.e.,

$$\mathbf{m}_{PM}(a)=.5,\ \mathbf{m}_{PM}(\sim a)=0.3,\ \text{and}\ \mathbf{m}_{PM}(\{a,\sim a\})=0.2.$$

malization procedure[8] in aggregating various items of evidence and thus yields simple analytical formulas.

Let us go back to our example of analytical procedures discussed above and express the auditor's judgment about the level of support obtained (or to be obtained when planning the audit) from the procedures for account 'A' in terms of algebraic expressions:

$$\mathbf{m}_{PA}(a)=1-APR_A, \tag{1}$$

$$\mathbf{m}_{PA}(\sim a)=0, \tag{2}$$

and

$$\mathbf{m}_{PA}(\{a,\sim a\})=APR_A, \tag{3}$$

where APR_A represents a number. Equation (1) implies that the analytical procedures performed by the auditor for account 'A' provide assurance that the account is not materially misstated with $(1-APR_A)$ degree of support.

Belief Functions and Plausibility Functions

In general, the total belief in a subset B of the frame Θ is given by:

$$\mathbf{Bel}(B)=\sum_{X\subseteq B}\mathbf{m}(X), \tag{4}$$

where X represents a set of elements of Θ, and the plausibility of B is given by:

$$\mathbf{PL}(B)=\sum_{B\cap X\neq\emptyset}\mathbf{m}(X)=1-\mathbf{Bel}(\sim B). \tag{5}$$

Intuitively, the plausibility of B is the degree to which B is plausible in the light of the evidence—the degree to which we do not disbelieve B or assign belief to its negation $\sim B$. Complete ignorance or lack of opinion about B is represented by $\mathbf{Bel}(B)=0$ and $\mathbf{PL}(B)=1$.

Consider again the numerical example discussed above. We have $\mathbf{m}_{PA}(a)=0.6$, $\mathbf{m}_{PA}(\sim a)=0$, and $\mathbf{m}_{PA}(a,\sim a)=0.4$. From equations (4) and (5), we obtain:

$$\mathbf{Bel}_{PA}(a)=\mathbf{m}_{PA}(a)=0.6,$$

$$\mathbf{Bel}_{PA}(\sim a)=\mathbf{m}_{PA}(\sim a)=0,$$

$$\mathbf{Bel}_{PA}(\{a,\sim a\})=\mathbf{m}_{PA}(a)+\mathbf{m}_{PA}(\sim a)+\mathbf{m}_{PA}(a,\sim a)=0.6+0+0.4=1.0,$$

---

[8] Consider two independent items of evidence with $\mathbf{m}_1$ and $\mathbf{m}_2$ representing the $\mathbf{m}$-values on a frame Θ. By Dempster's rule (Shafer 1976), the combined $\mathbf{m}$-value for a subset A of frame Θ is:

$$\mathbf{m}(A)=K^{-1}\Sigma\{\mathbf{m}_1(B_1)\mathbf{m}_2(B_2)\mid B_1\cap B_2=A, A\neq\emptyset\},$$

where K is the renormalization constant;

$$K=1-\Sigma\{\mathbf{m}_1(B_1)\mathbf{m}_2(B_2)\mid B_1\cap B_2=\emptyset\}.$$

The second term in K represents the conflict between the two items of evidence. If the conflict term is 1, i.e., if the two items of evidence exactly contradict each other, then $K=0$ and, in such a situation, the two items of evidence are not combinable. In other words, Dempster's rule cannot be used when $K=0$.

and

$$\boldsymbol{PL}_{PA}(a)=1-\boldsymbol{Bel}_{PA}(\sim a)=1,$$

$$\boldsymbol{PL}_{PA}(\sim a)=1-\boldsymbol{Bel}_{PA}(a)=1-.06=0.4.$$

The intuitive meaning of $\boldsymbol{Bel}_{PA}(a)=0.6$ is that the auditor has direct evidence from analytical procedures relevant to account '*A*' that '*a*' is true with 0.6 degree of support (i.e., the account is not materially misstated with degree 0.6). $\boldsymbol{Bel}_{PA}(\sim a)=0$ means that the auditor has no evidence from analytical procedures that the account is materially misstated (i.e., $\sim a$ is true).

Let us now consider $\boldsymbol{PL}_{PA}(a)=1$. What does it mean? We know that analytical procedures provide no belief to $\sim a(\boldsymbol{Bel}_{PA}(\sim a)=0)$. Since there is no support committed to just $\sim a$ all the probability mass could be assigned to *a*, which implies that $\boldsymbol{PL}_{PA}(a)=1$. Similarly, since $\boldsymbol{Bel}_{PA}(a)=0.6$ (i.e., 0.6 degree of belief is directly committed to *a*), the remaining amount 0.4 of uncommitted probability mass could be assigned to $\sim a$; that is, $\boldsymbol{PL}_{PA}(\sim a)=0.4$.

Going back to the **m**-values in equations (1)–(3), we obtain the following beliefs and plausibilities:

$$\boldsymbol{Bel}_{PA}(a)=1-APR_A,\ \boldsymbol{Bel}_{PA}(\sim a)=0,\text{ and }\boldsymbol{Bel}_{PA}(\{a,\sim a\})=1, \tag{6}$$

and

$$\boldsymbol{PL}_{PA}(a)=1,\text{ and }\boldsymbol{PL}_{PA}(\sim a)=APR_A. \tag{7}$$

The plausibility function, $\boldsymbol{PL}_{PA}(\sim a)=APR_A$, has an important interpretation. It provides a non-frequentist interpretation of the auditing concept of risk. This is a measure of how risky we feel it would be to stop with this evidence. According to the analytical procedures performed at the account level, we have $(1-APR_A)$ degree of belief that *a* is true, leaving a plausibility of $\boldsymbol{PL}_{PA}(\sim a)=APR_A$ that the account is materially misstated. This is the audit risk associated with the analytical procedures performed at the account level (see table 1 for definitions of other risks).

This plausibility interpretation of audit risk is conceptually in agreement with the thought process of the auditor when planning an audit. For example, if the auditor plans an audit of an account to obtain an overall assurance of 0.95 (i.e., $\boldsymbol{Bel}(a)=0.95$) that the account is not materially misstated then, in plausibility terms, it means that the auditor is planning the audit at the 0.05 level of plausibility for material error in the account (i.e., $\boldsymbol{PL}(\sim a)=0.05$). In other words, if the auditor had to stop after obtaining 0.95 level of assurance that '*a*' was true, then the evidence gathered up to that point would suggest that '$\sim a$' is plausible with degree 0.05; that is, there is a *maximum* risk of 0.05 that the account is materially misstated.[9]

In general, $\boldsymbol{Bel}(B)\leq\boldsymbol{PL}(B)$ for every subset *B* of our frame Θ. If we believe *B*, then we think *B* is plausible, but the converse is not necessarily true. A zero plausibility for a proposition means that we are sure that it is false (like a zero probability in the Bayesian

---

[9] Here is another example. Suppose the evidence gathered up to this point gives us the following **m**-values: $\mathbf{m}(a)=0.95$, $\mathbf{m}(\sim a)=0.02$, $\mathbf{m}(\{a,\sim a\})=0.03$. From equations (4) and (5), we obtain $\boldsymbol{Bel}(a)=0.95$, $\boldsymbol{PL}(a)=0.98$, $\boldsymbol{Bel}(\sim a)=0.02$, $\boldsymbol{PL}(\sim a)=0.05$. This means that we have 0.95 degree of belief that '*a*' is true and 0.02 degree of belief that '$\sim a$' is true. However, the plausibility of '$\sim a$' is 0.05, which means that if we had to stop at this point, we would be taking a maximum risk of 0.05 that '$\sim a$' is true although the belief in '$\sim a$' is only 0.02.

Table 1
List of Symbols and Their Definitions

Propositions

a – Account 'A' is not materially misstated.
$\sim a$ – Account 'A' is materially misstated.
f – The financial statement, F, is not materially misstated.
$\sim f$ – The financial statement, F, is materially misstated.
ao – There is no material misstatement related to objective 'O' of account 'A'.
$\sim ao$ – There is material misstatement related to objective 'O' of account 'A'.

m-Functions, Belief Functions, and Plausibility Functions

$\mathbf{m}_A(\bullet)$ – **m**-values at the level of account A for the proposition(s) in the argument.
$\mathbf{m}_F(\bullet)$ – **m**-values at the level of the financial statement for the proposition(s) in the argument.
$\mathbf{m}_{AO}(\bullet)$ – **m**-values at the level of the audit objective O of account A for the proposition(s) in the argument.
$\mathbf{m}_{CO}(\bullet)$ – **m**-values obtained from internal *controls* and accounting systems at the audit *objective* level for the proposition(s) in the argument.
$\mathbf{m}_{DO}(\bullet)$ – **m**-values obtained from *detailed* test of balance at the audit *objective* level for the proposition(s) in the argument.
$\mathbf{m}_{IA}(\bullet)$ – m-values obtained from *inherent* factors at the *account* level for the proposition(s) in the argument.
$\mathbf{m}_{IO}(\bullet)$ – m-values obtained from *inherent* factors at the audit *objective* level for the proposition(s) in the argument.
$\mathbf{m}_{IF}(\bullet)$ – m-values obtained from *inherent* factors at the *financial* statement level for the proposition(s) in the argument.
$\mathbf{m}_{PA}(\bullet)$ – **m**-values obtained from analytical *procedures* performed at the *account* level for the proposition(s) in the argument.
$\mathbf{m}_{PF}(\bullet)$ – **m**-values obtained from analytical *procedures* performed at the *financial statement* level for the proposition(s) in the argument.
$\mathbf{m}_{PO}(\bullet)$ – **m**-values obtained from analytical *procedures* performed at the audit *objective* level for the proposition(s) in the argument.
$\mathbf{Bel}_x(\bullet)$ – A belief function, x represents various indices as described above in the case of **m**.
$\mathbf{PL}_x(\bullet)$ – A plausibility function, x again represents various indices as described in the case of **m**.

Plausibility Functions for Material Misstatements
(i.e., Audit Risks in Belief-Function Framework)*

IR_{AO} – Plausibility of material misstatement in objective 'O' of account 'A' obtained from inherent factors at 'O' of 'A', or the risk (in the belief-function framework) associated with inherent factors at 'O' of account 'A'.
CR_{AO} – Control risk (plausibility of material misstatement on the basis of internal controls and accounting systems) for account 'A' at audit objective 'O'.
APR_{AO} – The risk associated with analytical procedures performed at the audit objective level.
DR_{AO} – The risk associated with detailed test of balance for account 'A' at objective 'O'.

$AR_{AO} = IR_{AO}APR_{AO}CR_{AO}DR_{AO}$, audit risk at the audit *objective* level when considering only the items of evidence that directly support the audit objective 'O' of account 'A'.

IR_A – The risk associated with inherent factors at the *account* level.
APR_A – The risk associated with analytical procedures performed at the *account* level.

IR_F – The risk associated with inherent factors at the *financial* statement level.
APR_F – The risk associated with analytical procedures performed at the *financial* statement level.

$AR'_{AO} = IR_F APR_F IR_A APR_A AR_{AO}$, total audit risk at the *objective* level considering all the items of evidence.

$AR'_A = IR_F APR_F IR_A APR_A \left[1 - \prod_O (1 - AR_{AO})\right]$, total audit risk at the *account* level considering all the items of evidence.

$AR'_F = IR_F APR_F \left[1 - \prod_A (1 - AR_A)\right]$, total audit risk at the *financial statement* level.

---

* Note that we have used the term "risk" for plausibility of material misstatement in the table and also in the text.

Table 2
m-Values, Belief, and Plausibility Functions for Different Audit Evidence Given in Figure 1

Evidence	*m-values*	*Belief*	*Plausibility*
Audit Objective Level:			
Inherent factors	$m_{IO}(ao)=1-IR_{AO}$ $m_{IO}(\sim ao)=0$ $m_{IO}(\{ao,\sim ao\})=IR_{AO}$	$Bel_{IO}(ao)=1-IR_{AO}$ $Bel_{IO}(\sim ao)=0$ $Bel_{IO}(\{ao,\sim ao\})=1$	$PL_{IO}(ao)=1$ $PL_{IO}(\sim ao)=IR_{AO}$
Analytical procedures	$m_{PO}(ao)=1-APR_{AO}$ $m_{PO}(\sim ao)=0$ $m_{PO}(\{ao,\sim ao\})=APR_{AO}$	$Bel_{PO}(ao)=1-APR_{AO}$ $Bel_{PO}(\sim ao)=0$ $Bel_{PO}(\{ao,\sim ao\})=1$	$PL_{PO}(ao)=1$ $PL_{PO}(\sim ao)=APR_{AO}$
Control factors and accounting systems	$m_{CO}(ao)=1-CR_{AO}$ $m_{CO}(\sim ao)=0$ $m_{CO}(\{ao,\sim ao\})=CR_{AO}$	$Bel_{CO}(ao)=1-CR_{AO}$ $Bel_{CO}(\sim ao)=0$ $Bel_{CO}(\{ao,\sim ao\})=1$	$PL_{CO}(ao)=1$ $PL_{CO}(\sim ao)=CR_{AO}$
Detailed tests of balance	$m_{DO}(ao)=1-DR_{AO}$ $m_{DO}(\sim ao)=0$ $m_{DO}(\{ao,\sim ao\})=DR_{AO}$	$Bel_{DO}(ao)=1-DR_{AO}$ $Bel_{DO}(\sim ao)=0$ $Bel_{DO}(\{ao,\sim ao\})=1$	$PL_{DO}(ao)=1$ $PL_{DO}(\sim ao)=DR_{AO}$
Account Level:			
Inherent factors	$m_{IA}(a)=1-IR_{A}$ $m_{IA}(\sim a)=0$ $m_{IA}(\{a,\sim a\})=IR_{A}$	$Bel_{IA}(a)=1-IR_{A}$ $Bel_{IA}(\sim a)=0$ $Bel_{IA}(\{a,\sim a\})=1$	$PL_{IA}(a)=1$ $PL_{IA}(\sim a)=IR_{A}$
Analytical procedures	$m_{PA}(a)=1-APR_{A}$ $m_{PA}(\sim a)=0$ $m_{PA}(\{a,\sim a\})=APR_{A}$	$Bel_{PA}(a)=1-APR_{A}$ $Bel_{PA}(\sim a)=0$ $Bel_{PA}(\{a,\sim a\})=1$	$PL_{PA}(a)=1$ $PL_{PA}(\sim a)=APR_{A}$
Financial Statement Level:			
Inherent factors	$m_{IF}(f)=1-IR_{F}$ $m_{IF}(\sim f)=0$ $m_{IF}(\{f,\sim f\})=IR_{F}$	$Bel_{IF}(f)=1-IR_{F}$ $Bel_{IF}(\sim f)=0$ $Bel_{IF}(\{f,\sim f\})=1$	$PL_{IF}(f)=1$ $PL_{IF}(\sim f)=IR_{F}$
Analytical procedures	$m_{PF}(f)=1-APR_{F}$ $m_{PF}(\sim f)=0$ $m_{PF}(\{f,\sim f\})=APR_{F}$	$Bel_{PF}(f)=1-APR_{F}$ $Bel_{PF}(\sim f)=0$ $Bel_{PF}(\{f,\sim f\})=1$	$PL_{PF}(f)=1$ $PL_{PF}(\sim f)=APR_{F}$

Note: The symbols are defined in table 1. Each item of audit evidence is assumed to be affirmative.

theory), but a zero degree of belief for a proposition means only that we see no reason to believe the proposition.

Similar explanations can be given to the m-values, belief functions, and plausibility functions for the other seven items of evidence presented in table 2. The individual m-values given in table 2 combined with Dempster's rule (see fn. 8) will give us the overall belief that the financial statement is fairly presented. Since these m-values are defined at different nodes in an evidential network (e.g., see fig. 1), combining them becomes a problem of propagating m-values (or belief functions) through the network. We have discussed this problem in a working paper[10] and have summarized the results in appendix A.

---

[10] This paper is available on request from the authors.

IV. m-Values Directly Defined at Each Node

To combine all the evidence in figure 1, we need to combine the **m**-values given in table 2. There are eight sets of **m**-values (four at the audit objective level, two at the account level, and two at the financial statement level; see table 2). Combining these **m**-values, in general, is very complex. However, the process is simplified if we proceed in two steps. First, we combine the **m**-values directly bearing on each node in figure 1. For example, we combine the **m**-values obtained from the four items of evidence at the audit objective level. Similarly, the total **m**-values bearing directly at the account level and the financial statement level will be the combination of two **m**-values defined at each level. Second, we propagate the above **m**-values obtained at each node using the results of appendix A. In the following paragraphs, we discuss the first step and provide analytical formulas for the resultant **m**-values at each level.

m-Values at the Financial Statement Level

We assume that the two items of evidence directly bearing at the financial statement level are: (1) inherent factors and (2) analytical procedures (see fig. 1). We further assume that these items of evidence are affirmative in nature. The corresponding **m**-values are given in table 2. Since there is no conflict among the evidence, we obtain the following values directly defined at the financial statement level using Dempster's rule (see fn. 8):

$$\begin{aligned}\mathbf{m}_F(f)&=\mathbf{m}_{IF}(f)\mathbf{m}_{PF}(f)+\mathbf{m}_{IF}(f)\mathbf{m}_{PF}(\{f,\sim f\})+\mathbf{m}_{IF}(\{f,\sim f\})\mathbf{m}_{PF}(f)\\&=(1-IR_F)(1-APR_F)+(1-IR_F)APR_F+IR_F(1-APR_F)\\&=1-IR_FAPR_F,\end{aligned}\tag{8}$$

$$\begin{aligned}\mathbf{m}_F(\sim f)=&\,\mathbf{m}_{IF}(\sim f)\mathbf{m}_{PF}(\sim f)+\mathbf{m}_{IF}(\sim f)\mathbf{m}_{PF}(\{f,\sim f\})\\&+\mathbf{m}_{IF}(\{f,\sim f\})\mathbf{m}_{PF}(\sim f)=0,\end{aligned}\tag{9}$$

and

$$\mathbf{m}_F(\{f,\sim f\})=\mathbf{m}_{IF}(\{f,\sim f\})\mathbf{m}_{PF}(\{f,\sim f\})=IR_FAPR_F.\tag{10}$$

From the above results, one can generalize that if the frame consists of only two elements, such as f, and $\sim f$, and the **m**-values for $\sim f$ are zero for all the evidence, then the combined **m**-value for $\sim f$ will be zero, irrespective of the number of items of evidence, and the combined **m**-value for the entire frame $\{f,\sim f\}$ will be the product of its **m**-values from each item of evidence.

m-Values at the Account Level

As shown in figure 1, we consider two items of evidence at the account level: (1) inherent factors and (2) analytical procedures. Again, it is assumed that these items of evidence are affirmative in nature, and the corresponding **m**-values are given in table 2. The combined **m**-values are similar to those in equations (8) through (10):

$$\mathbf{m}_A(a)=1-IR_AAPR_A,\tag{11}$$

$$\mathbf{m}_A(\sim a)=0,\tag{12}$$

and

$$\mathbf{m}_A(\{a, \sim a\}) = IR_A APR_A. \tag{13}$$

m-Values at the Audit Objective Level

In general, there are four items of evidence at the audit objective level (see fig. 1). Again, assume that they are all affirmative in nature. Then the corresponding **m**-values can be given as in table 2. We want to combine all the **m**-values obtained from these items of evidence. Since the **m**-values for $\sim ao$ for all the evidence are assuumed to be zero, the combined **m**-value for $\sim ao$ is zero (see eq. [9]), that is,

$$\mathbf{m}_{AO}(\sim ao) = 0. \tag{14}$$

Also, as seen in equation (10), since there is no conflict (i.e., $K=1$ in Dempster's rule: see fn. 8), the combined **m**-value for the entire frame $\{ao, \sim ao\}$ is equal to the product of the **m**-values for the frame from each item of evidence, that is,

$$\begin{aligned}\mathbf{m}_{AO}(\{ao, \sim ao\}) &= \mathbf{m}_{IO}(\{ao, \sim ao\})\mathbf{m}_{PO}(\{ao, \sim ao\})\mathbf{m}_{CO}(\{ao, \sim ao\})\mathbf{m}_{DO}(\{ao, \sim ao\}) \\ &= IR_{AO}APR_{AO}CR_{AO}DR_{AO}.\end{aligned} \tag{15}$$

Let us define a new term, AR_{AO}, for convenience as:

$$AR_{AO} = IR_{AO}APR_{AO}CR_{AO}DR_{AO}. \tag{16}$$

From equations (14) through (16) and the definition of **m** function, we obtain:

$$\begin{aligned}\mathbf{m}_{AO}(ao) &= 1 - \mathbf{m}_{AO}(\sim ao) - \mathbf{m}_{AO}(\{ao, \sim ao\}) \\ &= 1 - AR_{AO}.\end{aligned} \tag{17}$$

The corresponding belief and plausibility functions are (from eqs. [4] and [5]):

$$\mathbf{Bel}_{AO}(ao) = 1 - AR_{AO},$$

$$\mathbf{Bel}_{AO}(\sim ao) = 0, \text{ and } \mathbf{Bel}_{AO}(\{ao, \sim ao\}) = 1,$$

$$\mathbf{PL}_{AO}(ao) = 1,$$

and

$$\mathbf{PL}_{AO}(\sim ao) = AR_{AO}. \tag{18}$$

Thus, AR_{AO} is the total plausibility for $\sim ao$ at the audit objective level.

It is interesting to note from equations (16) and (18) that the total plausibility at the audit objective level for $\sim ao$ is the product of the individual plausibilities for $\sim ao$ (see table 2). As discussed earlier in section III, plausibility is one interpretation of audit risk. So equation (18) along with equation (16) represents the audit risk model at the audit objective level without considering any other evidence at the account level or the financial statement level. This formula is similar to that of SAS No. 47. However, it is incomplete as an overall model for the audit risk because it does not include evidence at the other levels.

In the remainder of this article, we develop analytical formulas for the overall plausibility of material misstatement in the financial statement and the account and compare and contrast them with the SAS No. 47 formula. Since we have now determined the m-values directly defined at each node of figure 1, we can use the results of appendix A to derive our formulas (see appendix B for details).

V. Audit-Risk Formulas in the Belief-Function Framework

In this section, we give formulas for the overall plausibility of material misstatement at various levels of the financial statement. These formulas have been derived in appendix B.

Total Audit Risk (Plausibility of Material Misstatement) at the Financial Statement Level

We have the following expression for total plausibility of material misstatement at the financial statement level (eq. B-7):

$$PL^t_F(\sim f)=AR^t_F=IR_F APR_F\left[1-\prod_A(1-AR_A)\right], \tag{19}$$

where AR_A and AR_{AO} are defined in equations (B-4) and (16), respectively, as:

$$AR_A=IR_A APR_A\left[1-\prod_O(1-AR_{AO})\right], \tag{20}$$

and

$$AR_{AO}=IR_{AO}APR_{AO}CR_{AO}DR_{AO}. \tag{21}$$

Also, the total belief that the financial statement is fairly presented is given by equation (B-7):

$$Bel^t_F(f)=1-IR_F APR_F\left[1-\prod_A(1-AR_A)\right]. \tag{22}$$

Equation (19) represents total plausibility of material misstatement in the financial statement or total audit risk, AR^t_F, at the financial statement level. The total belief that the financial statement is not materially misstated is given by equation (22). It should be noted that the algebraic form of equation (19) is very different from the formula discussed in SAS No. 47 or the Bayesian model. Unlike the audit risk model of SAS No. 47 or the Bayesian model, equation (19) takes into consideration all the evidence at all the levels of the financial statement. It also differs from SAS No. 47, of course, in the interpretation. Here, we interpret audit risk as a plausibility, not as a probability.

Total Audit Risk (Plausibility of Material Misstatement) at the Account Level

From equation (B-11), we have the following expression for total plausibility of material misstatement at the account level:

$$PL^t_A(\sim a)=AR^t_A=IR_F APR_F IR_A APR_A\left[1-\prod_O(1-AR_{AO})\right], \tag{23}$$

and the total belief that the account 'A' is not materially misstated as:

$$Bel^t_A(a)=1-AR^t_A=1-IR_F APR_F IR_A APR_A\left[1-\prod_O(1-AR_{AO})\right]. \tag{24}$$

Here, equation (23) represents total plausibility of material misstatement or total audit risk, AR^t_A, at the account level. The total belief or assurance that the account is not materially misstated is given by equation (24). It is again the result of aggregating all the evidence at the account level, whether the evidence is coming from the audit objective level, the financial statement level, or directly bearing on the account. It again differs from the Bayesian or SAS No. 47 formula. In equation (23), we find that AR'_A is the product of three types of plausibilities: (1) plausibility arising from inherent factors (i.e., the inherent risk, $IR_F IR_A$), (2) plausibility arising from analytical procedures (i.e., the analytical procedure risk, $APR_F APR_A$), and (3) plausibility arising from the evidence at the audit objective level for the account (i.e., the combined audit risk, $[1-\Pi_O(1-AR_{AO})]$). The third term represents 1 minus the level of support obtained from the procedures performed at the audit objective level. If no procedures are performed at that level, which means $AR_{AO}=1$, then the support is zero and the third term equals 1. We will give a numerical example later.

Total Audit Risk (Plausibility of Material Misstatement) at the Audit Objective Level

The total plausibility of material misstatement and total belief at the audit objective level are given by equations (B-15) and (B-17):

$$\mathbf{PL'_{AO}}(\sim ao)=AR'_{AO}=IR_F APR_F IR_A APR_A AR_{AO}; \tag{25}$$

$$\mathbf{Bel'_{AO}}(ao)=1-IR_F APR_F IR_A APR_A AR_{AO}. \tag{26}$$

Equation (25) represents the total plausibility that the audit objective '*AO*' will not be met when all the evidence at various levels has been aggregated. The total belief that the objective will be met is given by equation (26). As seen in equation (25), the total risk at the audit objective level is the product of three terms, $(IR_F APR_F)$, $(IR_A APR_A)$, and $(IR_{AO} APR_{AO} CR_{AO} DR_{AO})$, each defined at different levels. This formula resembles the multiplicative formula of SAS No. 47 if we separate the risks associated with inherent factors and analytical prodcedures:

$$AR'_{AO}=(IR_F IR_A IR_{AO})(APR_F APR_A APR_{AO})(CR_{AO} DR_{AO}). \tag{27}$$

The first factor in equation (27) determines the overall risk associated with inherent factors. Similarly, the second term represents the overall risk associated with analytical procedures performed at all levels. The third term is the product of control risk and detection risk. We must repeat that, although equation (27) is similar to the SAS No. 47 model, our interpretation of the risk is very different.

Numerical Example

Suppose we have only five accounts on the balance sheet and each account has five objectives. Suppose the auditor has gathered and evaluated all the relevant inherent factors at the level of the financial statement and the account and has assigned the following values for the respective plausibilities of material misstatement or risks: $IR_F=0.7$ and $IR_A=0.6$ for all the accounts. Also assume that the auditor has performed analytical procedures for various accounts and assigned a plausibility of material misstatement or risk of 0.4 to these procedures, but has not performed any analytical procedures at the financial statement level. Thus, $APR_A=0.4$ for all the accounts and $APR_F=1$. These values result in a total plausibility of error or risk at the financial statement level of 0.52, that is, $AR'_F=0.52$ (table 3). This implies that, on the basis of evidence accumulated at the financial statement level and the account level, the auditor finds that the financial statement is not materially misstated with a belief or assurance

Table 3
Audit Risk Model in Belief-Function Framework with IR_A, APR_A, and IR_F as Inputs

Panel A. Input Risks:

(1) *Risks at the Audit Objective Level*

	IR_{AO}					APR_{AO}					CR_{AO}					DR_{AO}				
	Objective					*Objective*					*Objective*					*Objective*				
Account	O_1	O_2	O_3	O_4	O_5	O_1	O_2	O_3	O_4	O_5	O_1	O_2	O_3	O_4	O_5	O_1	O_2	O_3	O_4	O_5
A_1	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
A_2	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
A_3	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
A_4	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
A_5	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0

(2) *Risks at the Account Level*

Account	IR_A	APR_A
A_1	0.6	0.4
A_2	0.6	0.4
A_3	0.6	0.4
A_4	0.6	0.4
A_5	0.6	0.4

(3) *Risks at the Financial Statement Level*

$IR_F = 0.7$

$APR_F = 1.0$

Panel B. Output Risks:

	(1) *Total Risk at the Audit Objective Level* (AR^t_{AO}) (*From eq.* [25])					(2) *Total Risk at the Account Level* (AR^t_A) (*From eq.* [23])
Account	O_1	O_2	O_3	O_4	O_5	
A_1	0.168	0.168	0.168	0.168	0.168	0.168
A_2	0.168	0.168	0.168	0.168	0.168	0.168
A_3	0.168	0.168	0.168	0.168	0.168	0.168
A_4	0.168	0.168	0.168	0.168	0.168	0.168
A_5	0.168	0.168	0.168	0.168	0.168	0.168

(3)
Total Risk and Belief at the Financial Statement Level
(*From eq.* [19])

$AR^t_F = 0.52$
$Bel^t_F(f) = 0.48$

Note: The term "risk" is used for plausibility of material misstatement. The symbols used in the table are defined in table 1.

of 0.48. This information would help the auditor plan a more efficient audit than is possible with the SAS No. 47 approach.

Let us now assume that the auditor plans to proceed to the detailed level and considers the following steps: (1) collect and evaluate inherent factors at the audit objective level, (2) perform analytical procedures at the audit objective level if appropriate, (3) study and evaluate client's accounting systems and control procedures and perform test of transactions, and (4) perform direct test of balance. The results of the auditor's judgment are shown in terms of plausibilities of material misstatement or risks as inputs in tables 4 through 7.

In table 4, we see that there is almost no impact of inherent factors at the audit objective level on AR'_F and AR'_A. However, AR'_{AO} reduces from 16.8 percent to 11.8 percent. Consideration of analytical procedures at the audit objective level reduces AR'_F to 50 percent, AR'_A to 15.7 percent, and AR'_{AO} to 7.1 percent, as shown in table 5. When accounting systems and control procedures are included in the model, the total plausibility of material misstatement (i.e., the total risk at various levels) is further reduced. $AR'_F = 28$ percent, AR'_A varies between 6.0 and 10.1 percent, and AR'_{AO} between 1.4 and 2.8 percent (see table 6).

At this stage, the auditor decides about the extent, timing, and nature of the detailed test of balance so as to obtain the overall plausibility of material misstatement at 0.05 or an overall assurance of 0.95. As shown in table 7, for the level of risk given in the table for each audit objective from the corresponding detailed test of balance, we obtain the desired 0.95 level of assurance or belief that the financial statement is not materially misstated (i.e., an overall plausibility or total audit risk of 0.05). The risks at the other levels vary as follows: AR'_A varies between 1.0 and 1.4 percent, and AR'_{AO} varies between 0.2 and 0.3 percent (see table 7).

It is important to note that consideration of the structure of evidence in our plausibility models makes the audit process more efficient. In other words, the auditor will plan less extensive tests at the audit objective level when the evidence at the financial statement level and account level is positive. One should keep in mind that certain required procedures at the detailed level must be performed; the evidence at the financial and account levels would affect only the extent of testing. As seen from table 8, when the evidence at the account level and the financial statement level is not included in the plausibility model (as done in SAS No. 47 for the audit risk model), the total plausibility of material misstatement or the total audit risk at the financial statement level is high ($AR'_F = 29$ percent) or the total belief that the financial statement is not materially misstated is low (0.71). Thus, without an explicit treatment of the evidence at the financial statement level and the account level, the auditor will always underestimate the overall assurance and will collect more evidence than necessary at the detail level.

VI. Limitations of the Models

We must emphasize that there are strong limitations on the applicability of the formulas we have derived, limitations that are shared by the SAS No. 47 formula and existing Bayesian formulas.

First, since we have considered only binary variables, we have distinguished only whether an account is materially misstated or not. We have not distinguished between material misstatement due to overstatement or understatement. This limitation will

Table 4
Audit Risk Model in Belief-Function Framework with IR_{AO}, IR_A, APR_A, and IR_F as Inputs

Panel A. Input Risks:

(1) Risks at the Audit Objective Level

	IR_{AO}					APR_{AO}					CR_{AO}					DR_{AO}				
	Objective					Objective					Objective					Objective				
Account	O_1	O_2	O_3	O_4	O_5	O_1	O_2	O_3	O_4	O_5	O_1	O_2	O_3	O_4	O_5	O_1	O_2	O_3	O_4	O_5
A_1	0.7	0.7	0.7	0.7	0.7	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
A_2	0.7	0.7	0.7	0.7	0.7	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
A_3	0.7	0.7	0.7	0.7	0.7	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
A_4	0.7	0.7	0.7	0.7	0.7	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
A_5	0.7	0.7	0.7	0.7	0.7	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0

(2) Risks at the Account Level

Account	IR_A	APR_A
A_1	0.6	0.4
A_2	0.6	0.4
A_3	0.6	0.4
A_4	0.6	0.4
A_5	0.6	0.4

(3) Risks at the Financial Statement Level

$IR_F=0.7$

$APR_F=1.0$

Panel B. Output Risks:

	(1) *Total Risk at the Audit Objective Level* (AR'_{AO}) (From eq. [25])					(2) *Total Risk at the Account Level* (AR'_A) (From eq. [23])
Account	O_1	O_2	O_3	O_4	O_5	
A_1	0.118	0.118	0.118	0.118	0.118	0.168
A_2	0.118	0.118	0.118	0.118	0.118	0.168
A_3	0.118	0.118	0.118	0.118	0.118	0.168
A_4	0.118	0.118	0.118	0.118	0.118	0.168
A_5	0.118	0.118	0.118	0.118	0.118	0.168

(3)
Total Risk and Belief at the Financial Statement Level
(From eq. [19])

$AR'_F=0.52$
$Bel'_F(\bar{f})=0.48$

Note: The term "risk" is used for plausibility of material misstatement. The symbols used in the table are defined in table 1.

Table 5

Audit Risk Model in Belief-Function Framework with IR_{AO}, APR_{AO}, IR_A, APR_A, and IR_F as Inputs

Panel A. Input Risks:

(1) Risks at the Audit Objective Level

	IR_{AO}					APR_{AO}					CR_{AO}					DR_{AO}				
	Objective					*Objective*					*Objective*					*Objective*				
Account	O_1	O_2	O_3	O_4	O_5	O_1	O_2	O_3	O_4	O_5	O_1	O_2	O_3	O_4	O_5	O_1	O_2	O_3	O_4	O_5
A_1	0.7	0.7	0.7	0.7	0.7	0.6	0.6	0.6	0.6	0.6	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
A_2	0.7	0.7	0.7	0.7	0.7	0.6	0.6	0.6	0.6	0.6	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
A_3	0.7	0.7	0.7	0.7	0.7	0.6	0.6	0.6	0.6	0.6	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
A_4	0.7	0.7	0.7	0.7	0.7	0.6	0.6	0.6	0.6	0.6	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
A_5	0.7	0.7	0.7	0.7	0.7	0.6	0.6	0.6	0.6	0.6	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0

(2) Risks at the Account Level

Account	IR_A	APR_A
A_1	0.6	0.4
A_2	0.6	0.4
A_3	0.6	0.4
A_4	0.6	0.4
A_5	0.6	0.4

(3) Risks at the Financial Statement Level

$IR_F=0.7$

$APR_F=1.0$

Panel B. Output Risks:

	(1) *Total Risk at the Audit Objective Level* (AR'_{AO}) (*From eq.* [25])					(2) *Total Risk at the Account Level* (AR'_A) (*From eq.* [23])	(3) *Total Risk and Belief at the Financial Statement Level* (*From eq.* [19])
Account	O_1	O_2	O_3	O_4	O_5		
A_1	0.071	0.071	0.071	0.071	0.071	0.157	
A_2	0.071	0.071	0.071	0.071	0.071	0.157	$AR'_F=0.50$
A_3	0.071	0.071	0.071	0.071	0.071	0.157	$Bel'_F(f)=0.50$
A_4	0.071	0.071	0.071	0.071	0.071	0.157	
A_5	0.071	0.071	0.071	0.071	0.071	0.157	

Note: The term "risk" is used for plausibility of material misstatement. The symbols used in the table are defined in table 1.

Table 6

Audit Risk Model in Belief-Function Framework with IR_{AO}, APR_{AO}, CR_{AO}, IR_A, APR_A, and IR_F as Inputs

Panel A. Input Risks:

(1) Risks at the Audit Objective Level

	IR_{AO}					APR_{AO}					CR_{AO}					DR_{AO}				
	Objective					*Objective*					*Objective*					*Objective*				
Account	O_1	O_2	O_3	O_4	O_5	O_1	O_2	O_3	O_4	O_5	O_1	O_2	O_3	O_4	O_5	O_1	O_2	O_3	O_4	O_5
A_1	0.7	0.7	0.7	0.7	0.7	0.6	0.6	0.6	0.6	0.6	0.2	0.2	0.2	0.2	0.2	1.0	1.0	1.0	1.0	1.0
A_2	0.7	0.7	0.7	0.7	0.7	0.6	0.6	0.6	0.6	0.6	0.2	0.2	0.2	0.2	0.2	1.0	1.0	1.0	1.0	1.0
A_3	0.7	0.7	0.7	0.7	0.7	0.6	0.6	0.6	0.6	0.6	0.4	0.4	0.4	0.4	0.4	1.0	1.0	1.0	1.0	1.0
A_4	0.7	0.7	0.7	0.7	0.7	0.6	0.6	0.6	0.6	0.6	0.2	0.2	0.2	0.2	0.2	1.0	1.0	1.0	1.0	1.0
A_5	0.7	0.7	0.7	0.7	0.7	0.6	0.6	0.6	0.6	0.6	0.2	0.2	0.2	0.2	0.2	1.0	1.0	1.0	1.0	1.0

(2) Risks at the Account Level

Account	IR_A	APR_A
A_1	0.6	0.4
A_2	0.6	0.4
A_3	0.6	0.4
A_4	0.6	0.4
A_5	0.6	0.4

(3) Risks at the Financial Statement Level

$IR_F=0.7$

$APR_F=1.0$

Panel B. Output Risks:

	(1) Total Risk at the Audit Objective Level (AR^t_{AO}) (From eq. [25])					(2) Total Risk at the Account Level (AR^t_A) (From eq. [23])
Account	O_1	O_2	O_3	O_4	O_5	
A_1	0.071	0.014	0.014	0.014	0.014	0.060
A_2	0.071	0.014	0.014	0.014	0.014	0.060
A_3	0.071	0.028	0.028	0.028	0.028	0.101
A_4	0.071	0.014	0.014	0.014	0.014	0.060
A_5	0.071	0.014	0.014	0.014	0.014	0.060

(3)
Total Risk and Belief
at the Financial Statement Level
(From eq. [19])

$AR^t_F=0.28$
$\mathbf{Bel}^t_F(f)=0.72$

Note: The term "risk" is used for plausibility of material misstatement. The symbols used in the table are defined in table 1.

Table 7

Audit Risk Model in Belief-Function Framework with IR_{AO}, APR_{AO}, CR_{AO}, DR_{AO}, IR_A, APR_A, and IR_F as Inputs

Panel A. Input Risks:

(1) Risks at the Audit Objective Level

	IR_{AO}					APR_{AO}					CR_{AO}					DR_{AO}				
	Objective					Objective					Objective					Objective				
Account	O_1	O_2	O_3	O_4	O_5	O_1	O_2	O_3	O_4	O_5	O_1	O_2	O_3	O_4	O_5	O_1	O_2	O_3	O_4	O_5
A_1	0.7	0.7	0.7	0.7	0.7	0.6	0.6	0.6	0.6	0.6	0.2	0.2	0.2	0.2	0.2	0.15	0.15	0.15	0.15	0.15
A_2	0.7	0.7	0.7	0.7	0.7	0.6	0.6	0.6	0.6	0.6	0.2	0.2	0.2	0.2	0.2	0.15	0.15	0.15	0.15	0.15
A_3	0.7	0.7	0.7	0.7	0.7	0.6	0.6	0.6	0.6	0.6	0.4	0.4	0.4	0.4	0.4	0.10	0.10	0.10	0.10	0.10
A_4	0.7	0.7	0.7	0.7	0.7	0.6	0.6	0.6	0.6	0.6	0.2	0.2	0.2	0.2	0.2	0.15	0.15	0.15	0.15	0.15
A_5	0.7	0.7	0.7	0.7	0.7	0.6	0.6	0.6	0.6	0.6	0.2	0.2	0.2	0.2	0.2	0.15	0.15	0.15	0.15	0.15

(2) Risks at the Account Level

Account	IR_A	APR_A
A_1	0.6	0.4
A_2	0.6	0.4
A_3	0.6	0.4
A_4	0.6	0.4
A_5	0.6	0.4

(3) Risks at the Financial Statement Level

$IR_F = 0.7$

$APR_F = 1.0$

Panel B. Output Risks:

	(1) Total Risk at the Audit Objective Level (AR'_{AO}) (From eq. [25])					(2) Total Risk at the Account Level (AR'_A) (From eq. [23])	(3) Total Risk and Belief at the Financial Statement Level (From eq. [19])
Account	O_1	O_2	O_3	O_4	O_5		
A_1	0.002	0.002	0.002	0.002	0.002	0.010	
A_2	0.002	0.002	0.002	0.002	0.002	0.010	$AR'_F = 0.05$
A_3	0.003	0.003	0.003	0.003	0.003	0.014	$Bel'_F(f) = 0.95$
A_4	0.002	0.002	0.002	0.002	0.002	0.010	
A_5	0.002	0.002	0.002	0.002	0.002	0.010	

Note: The term "risk" is used for plausibility of material misstatement. The symbols used in the table are defined in table 1.

Table 8

Audit Risk Model in Belief-Function Framework with IR_{AO}, APR_{AO}, CR_{AO}, DR_{AO} as Inputs

Panel A. Input Risks:

(1) Risks at the Audit Objective Level

	IR_{AO}					APR_{AO}					CR_{AO}					DR_{AO}				
	Objective					*Objective*					*Objective*					*Objective*				
Account	O_1	O_2	O_3	O_4	O_5	O_1	O_2	O_3	O_4	O_5	O_1	O_2	O_3	O_4	O_5	O_1	O_2	O_3	O_4	O_5
A_1	0.7	0.7	0.7	0.7	0.7	0.6	0.6	0.6	0.6	0.6	0.2	0.2	0.2	0.2	0.2	0.15	0.15	0.15	0.15	0.15
A_2	0.7	0.7	0.7	0.7	0.7	0.6	0.6	0.6	0.6	0.6	0.2	0.2	0.2	0.2	0.2	0.15	0.15	0.15	0.15	0.15
A_3	0.7	0.7	0.7	0.7	0.7	0.6	0.6	0.6	0.6	0.6	0.4	0.4	0.4	0.4	0.4	0.10	0.10	0.10	0.10	0.10
A_4	0.7	0.7	0.7	0.7	0.7	0.6	0.6	0.6	0.6	0.6	0.2	0.2	0.2	0.2	0.2	0.15	0.15	0.15	0.15	0.15
A_5	0.7	0.7	0.7	0.7	0.7	0.6	0.6	0.6	0.6	0.6	0.2	0.2	0.2	0.2	0.2	0.15	0.15	0.15	0.15	0.15

(2) Risks at the Account Level

Account	IR_A	APR_A
A_1	1.0	1.0
A_2	1.0	1.0
A_3	1.0	1.0
A_4	1.0	1.0
A_5	1.0	1.0

(3) Risks at the Financial Statement Level

$IR_F = 1.0$

$APR_F = 1.0$

Panel B. Output Risks:

	(1) *Total Risk at the Audit Objective Level (AR'_{AO}) (From eq. [25])*					(2) *Total Risk at the Account Level (AR'_A) (From eq. [23])*	(3) *Total Risk and Belief at the Financial Statement Level (From eq. [19])*
Account	O_1	O_2	O_3	O_4	O_5		
A_1	0.013	0.013	0.013	0.013	0.013	0.061	
A_2	0.013	0.013	0.013	0.013	0.013	0.061	$AR'_F = 0.29$
A_3	0.017	0.017	0.017	0.017	0.017	0.081	$Bel'_F(f) = 0.71$
A_4	0.013	0.013	0.013	0.013	0.013	0.061	
A_5	0.013	0.013	0.013	0.013	0.013	0.061	

Note: The term "risk" is used for plausibility of material misstatement. The symbols used in the table are defined in table 1.

make the audit process less efficient. For example, if there are two accounts, one materially overstated and the other materially understated by the same amount, and the auditor feels that the combination of the two accounts is fairly stated because of the offsetting errors, the present approach will suggest that the combination is materially misstated and hence lead to inefficiency.

Second, we have not considered immaterial errors in individual accounts that might add up to a material error in the whole financial statement. This will also make the audit less efficient.

Third, we have assumed that each account or audit objective is equally important. This assumption may make the audit process less efficient because the auditor may still have to obtain a high level of assurance for an unimportant audit objective.

We have already mentioned other limitations due to our simplifying assumptions. We considered only a tree-type evidential structure, and only affirmative evidence is considered. As we have already explained, it may be feasible to derive formulas with these assumptions relaxed to some extent, but an algorithmic approach to the more complex case is probably preferable.

When discussing the increased efficiency possible with more accurate representation of the structure of the audit evidence, we must always bear in mind that the decrease in audit effort that is implied by plans based on such structure can decrease audit effectiveness if the inputs to the model cannot be estimated accurately by auditors who employ it. We think this gives another advantage to the belief-function approach over the Bayesian approach, since the Bayesian approach demands, in general, probability judgments that may not be available or meaningful.

VII. Conclusion

This article has used the structure of audit evidence currently assumed in auditing standards. We have represented this structure by a network of variables; including variables that represented the financial statement as a whole, various accounts, and various audit objectives of each account. For simplicity, we have considered only tree-type evidential structures and only variables with two possible values.

We have used the belief-function framework to relate audit risk to plausibility of material misstatement and have derived formulas[11] for the total plausibility at the three levels: financial statement level, account level, and audit objective level. The formula for the total plausibility of material misstatement at the audit objective level resembles the formula of SAS No. 47. The total plausibility of material misstatement associated with inherent factors is the product of risks at the financial statement level, the account level, and the audit objective level. Similarly, the plausibility of error associated with analytical procedures is the product of risks at the three levels.

The formulas at the other two levels are significantly different from the SAS No. 47 model. Unlike the SAS No. 47 model or the Bayesian model, all the formulas developed in this article aggregate all the evidence obtained from procedures performed at various levels of the account.

It must be emphasized that interpretation of audit risk in this article is significantly different from the probability interpretation used in the auditing literature. For example, in the belief-function framework, the detection risk, DR_{AO}, is the plausibility that objective 'O' of account 'A' is materially misstated, whereas the SAS No. 47 model interprets DR_{AO} as the probability that the auditor's procedures will fail to detect material misstatements related to objective 'O' of account 'A', given that the internal control procedures have failed to prevent or detect and correct such misstatements.

---

[11] As Boritz (1990) has shown, almost identical expressions for the overall audit risk can be obtained with probability theory.

The formulas derived here provide an audit planning tool that can be used to determine the level of assurance to be obtained from various sources to achieve a desired level of overall assurance or an overall plausibility of material misstatement. Unlike the SAS No. 47 formula, these formulas take into consideraion all the items of evidence gathered by the auditor, whether that evidence bears on the objective level, the account level, or the financial statement level. Thus, our approach provides an improvement over the SAS No. 47 model for planning an efficient audit and evaluating the results.

There are many other issues that need to be addressed, including (1) the level of testing needed for a desired level of support or belief, (2) the level of support or belief obtained from a given statistical result of a test procedure for different variables, (3) how to integrate belief aggregation with the cost of evidence gathering to obtain the most effective and efficient audit strategy, and (4) whether a belief-function approach to planning and evaluation of an audit is cost effective. All of these issues require further research.

Appendix A
Propagation of m-Values in an "And" Tree

Propagation of beliefs (or m-values) in a network is quite complex. Extensive work on this topic is reported in the artificial intelligence literature. To mention a few studies, the work by Shafer (1976, 1986), Shafer and Shenoy (1988, 1990), Shenoy and Shafer (1988, 1990), Zarley et al. (1988), and Shenoy (1989) is important for this purpose.

We will summarize results of propagation of beliefs in an "and" tree. These results will be useful in deriving analytical formulas in appendix B. There are two important directions for propagation for the overall aggregation of beliefs in an "and" tree. One direction is from the subobjectives to the main objective. The second direction is to propagate to a given subobjective from the main objective and the other subobjectives. The details of this work are given in a working paper that is available from the authors on request.

Propagation of m-Values from Audit Objectives to the Respective Accounts

Here, we summarize the results of propagation of m-values in an "and" tree (fig. 1) from subobjectives to the main objectives, that is, from audit objectives to corresponding accounts or from accounts to the financial statement. In the case of propagation from audit objectives (AOs) to an account 'A', one obtains the following values:

$$\mathbf{m}'_{A-\text{all }AO}(a)=\prod_{i=1}^{n}\mathbf{m}_{AO_i}(ao_i), \tag{A-1}$$

$$\begin{aligned}\mathbf{m}'_{A-\text{all }AO}(\sim a)=&\sum_{i=1}^{n}\mathbf{m}_{AO_i}(\sim ao_i)\prod_{j\neq i}^{n}\mathbf{m}_{AO_j}(ao_j)\\&+\sum_{i=1}^{n}\sum_{j\neq i}^{n}\mathbf{m}_{AO_i}(\sim ao_i)\mathbf{m}_{AO_j}(\sim ao_j)\prod_{k\neq i,\,k\neq j}^{n}\mathbf{m}_{AO_k}(ao_k)\\&+\ldots+\ldots+\prod_{i=1}^{n}\mathbf{m}_{AO_i}(\sim ao_i)\\&+\sum_{i=1}^{n}\mathbf{m}_{AO_i}(\sim ao_i)\prod_{j\neq i}^{n}\mathbf{m}_{AO_j}(\{ao_j,\sim ao_j\})\end{aligned}$$

$$+\sum_{i=1}^{n}\sum_{j\neq i}^{n}\mathbf{m}_{AO_i}(\sim ao_i)\mathbf{m}_{AO_j}(\sim ao_j)\prod_{k\neq i,\,k\neq j}^{n}\mathbf{m}_{AO_k}(\{ao_k,\sim ao_k\})+\ldots+$$

$$+\sum_{i=1}^{n}\prod_{j\neq i}^{n}\mathbf{m}_{AO_i}(\sim ao_i)\mathbf{m}_{AO_j}(\{ao_j,\sim ao_j\}), \tag{A-2}$$

and

$$\mathbf{m}'_{A-\text{all }AO}(\{a,\sim a\})=1-\mathbf{m}'_{A-\text{all }AO}(a)-\mathbf{m}'_{A-\text{all }AO}(\sim a). \tag{A-3}$$

These values may look very complex but they are intuitive. For example, equation (A-1) represents the propagated **m**-value for the state 'a' that the account 'A' is not materially misstated. This value is equal to the product of all the **m**-values defined at each audit objective for 'ao' that the objective is met. This result is similar to the probability rule giving the probability that the account is not materially misstated as equal to the product of the individual probabilities that the audit objectives are met provided the "and" relationship is valid.

Equation (A-2) is also intuitive. It represents the resultant **m**-value for '$\sim a$' received from all its audit objectives. The account is materially misstated when at least one or all of the objectives are not met. There are several situations that contribute to this condition: (1) at least one of the objectives is not met but the rest have been met, (2) all the objectives are not met, and (3) at least one of the objectives is not met but we have no knowledge whether the rest have been met or not. The first two terms in equation (A-2) represent the first situation. The third term represents the second situation. The last three terms represent the third situation.

For affirmative items of evidence (i.e., for $\mathbf{m}_A(\sim a)=0$ and $\mathbf{m}_{AO_j}(\sim ao_j)=0$ for all j), equation (A-2) reduces to the following simple form:

$$\mathbf{m}'_{A-\text{all }AO}(\sim a)=0, \tag{A-4}$$

while equations (A-1) and (A-3) remain unchanged.

Propagation of **m***-Values to a Given Subobjective from the Main Objective and the Other Subobjectives*

In this section, we want to summarize the results of propagation to a given subobjective from the main objective and the other subobjectives. For example, the **m**-values propagated to a given audit objective AO_i from account 'A' and the remaining audit objectives can be given by:

$$\mathbf{m}'_{AO_i-A\text{ \& all other }AO\text{'s}}(ao_i)=K_i^{-1}\mathbf{m}_A(a)\prod_{j\neq i}^{n}[1-\mathbf{m}_{AO_j}(\sim ao_j)], \tag{A-5}$$

$$\mathbf{m}'_{AO_i-A\text{ \& all other }AO\text{'s}}(\sim ao_i)=K_i^{-1}\mathbf{m}_A(\sim a)\prod_{j\neq i}^{n}\mathbf{m}_{AO_j}(ao_j), \tag{A-6}$$

$$\begin{aligned}\mathbf{m}'_{AO_i-A\text{ \& all other }AO\text{'s}}(\{ao_i,\sim ao_i\})&=1-\mathbf{m}'_{AO_i-A\text{ \& all other }AO\text{'s}}(ao_i)\\&\quad-\mathbf{m}'_{AO_i-A\text{ \& all other }AO\text{'s}}(\sim ao_i),\end{aligned} \tag{A-7}$$

where K_i is the renormalization constant that can be written as $K_i=[1-\mathbf{m}_A(a)C_i]$ and C_i is given by:

$$\begin{aligned}C_i=&\sum_{j\neq i}^{n}\mathbf{m}_{AO_j}(\sim ao_j)\prod_{k\neq j}^{n}\mathbf{m}_{AO_k}(ao_k)\\
&+\sum_{j\neq i}^{n}\sum_{k\neq j}^{n}\mathbf{m}_{AO_j}(\sim ao_j)\mathbf{m}_{AO_k}(\sim ao_k)\prod_{l\neq j,\,l\neq k}^{n}\mathbf{m}_{AO_l}(ao_l)\\
&+\ldots+\ldots+\prod_{j\neq i}^{n}\mathbf{m}_{AO_j}(\sim ao_j)\\
&+\sum_{j\neq i}^{n}\mathbf{m}_{AO_j}(\sim ao_j)\prod_{k\neq j}^{n}\mathbf{m}_{AO_k}(\{ao_k,\sim ao_k\})\\
&+\sum_{j\neq i}^{n}\sum_{k\neq i}^{n}\mathbf{m}_{AO_j}(\sim ao_j)\mathbf{m}_{AO_k}(\sim ao_k)\prod_{l\neq i,\,l\neq j}^{n}\mathbf{m}_{AO_l}(\{ao_l,\sim ao_l\})+\ldots+\\
&+\sum_{j\neq i}^{n}\prod_{k\neq j}^{n}\mathbf{m}_{AO_k}(\sim ao_k)\mathbf{m}_{AO_j}(\{ao_j,\sim ao_j\}).\end{aligned} \tag{A-8}$$

Again, the above equations may appear complex but the results are intuitive. For example, the **m**-value expressed in equation (A-5) represents that the audit objective AO_i is met if account 'A' is not materially misstated ($\mathbf{m}_A(a)\neq 0$) and the other audit objectives are either met or their status is unknown ($\mathbf{m}_{AO_j}(\sim ao_j)<1$). Equation (A-6) is easier to understand. It tells us that there is finite support for the audit objective's not having been met if the account is found to be materially misstated but the other audit objectives have been met. The renormalization constant, K_i, in equations (A-5) through (A-7) represents 1 minus the conflict term. In general, the conflict term is defined as the sum of all the products of **m**-values of the sets of elements whose intersection is a null set. In our case, conflict will occur when there is a finite support that the account is not materially misstated ($\mathbf{m}_A(a)\neq 0$) and at least one of the other audit objectives is not met ($C_i\neq 0$).

In the case of affirmative items of evidence (i.e., $\mathbf{m}_A(\sim a)=0$ and $\mathbf{m}_{AO_j}(\sim ao_j)=0$ for all j), which is the case assumed in the present work, the renormalization constant becomes 1 and equations (A-5) and (A-6) reduce to:

$$\mathbf{m}'_{AO_i\text{—}A\ \&\ \text{all other}\ AO\text{'s}}(ao_i)=\mathbf{m}_A(a), \tag{A-9}$$

$$\mathbf{m}'_{AO_i\text{—}A\ \&\ \text{all other}\ AO\text{'s}}(\sim ao_i)=0, \tag{A-10}$$

while equation (A-7) remains unchanged. From equation (A-9), we find that, in the case of affirmative items of evidence, when **m**-values propagate to a given subobjective from the main objective and the other subobjectives, the **m**-values received by the subobjective are not a function of the **m**-values of the other subobjectives.

Appendix B

Derivation of Plausibility Formulas or Audit Risk Formulas

In this appendix, our main objective is to derive formulas for plausibility of material misstatement (i.e., audit-risk formulas; see fn. 2) within the belief-function framework for an evidential network (an "and" tree) given in figure 1. To achieve this objective, we need to aggregate all the items of evidence in figure 1. As discussed earlier, the process of combining various **m**-values in a network is, in fact, a process of propagating **m**-values through the network of variables. We will use the general results presented in appendix A to derive the audit risk formulas in the following sections. As mentioned earlier, the following assumptions are made in the derivation of the formulas.

First, each node (variable, such as an account or audit objective) in the tree is assumed to be a binary variable. In fact, the variables in an "and" tree are binary by definition. The general approach of aggregating evidence by using the belief-function framework is still valid for nonbinary variables, but the analytical formulas for such cases will be intractable. Second, each item of evidence is assumed to be affirmative in nature (i.e., the evidence provides positive support to the assertion and no support to its negation).

Financial Statement Level

The objective here is to derive a formula for the total plausibility of material misstatement or audit risk, at the financial statement level. This is achieved by combining all **m**-values that are either defined directly at the financial statement node (F) or have been propagated to F from all other nodes in the network (see fig. 2). We will proceed with the propagation process in two steps. First, we will consider propagation from audit objectives (AO's) to an account (A) and then combine these values from the audit objectives with the **m**-values at the account. Second, we propagate the resultant **m**-values from various accounts to the financial statement node and combine them with the **m**-values at the financial statement level. This process will yield the desired audit risk (plausibility of material misstatement) formula.

For the first step above, as an illustration, let us consider propagation of **m**-values defined at the three audit objectives of account 'A_3' in figure 2 to the account. These **m**-values are the result of combining all four items of evidence at the audit objective level and are given, in general terms, in equations (14) through (17). When these **m**-values are propagated to account A_3, we obtain the following values using equations (A-1), (A-3), and (A-4):

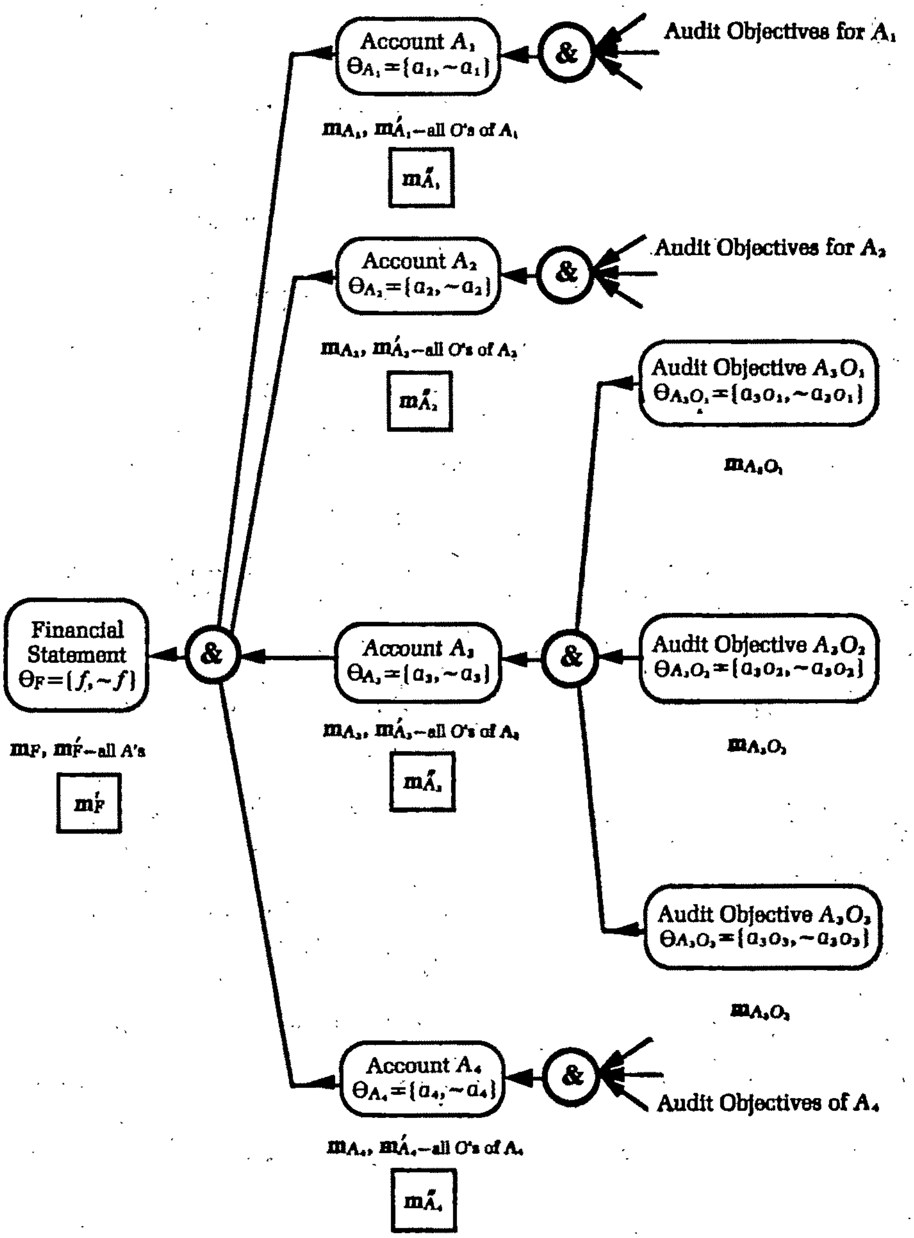
$$\mathbf{m}'_{A_3\text{—all}\ O\text{'s of}\ A_3}(a_3)=\prod_{i=1}^{3}\mathbf{m}_{A_3O_i}(a_3o_i)=\prod_{i=1}^{3}(1-AR_{A_3O_i}),$$

$$\mathbf{m}'_{A_3\text{—all}\ O\text{'s of}\ A_3}(\sim a_3)=0,$$

$$\mathbf{m}'_{A_3\text{—all}\ O\text{'s of}\ A_3}(\{a_3o_i,\sim a_3o_i\})=1-\prod_{i=1}^{3}(1-AR_{A_3O_i}), \tag{B-1}$$

where AR_{AO}'s are defined as in equation (16).

Figure 2
Propagation of m-Values from Audit Objectives and Accounts to the Financial Statement


Account A_1
$\Theta_{A_1}=\{a_1,\sim a_1\}$
&
Audit Objectives for A_1
m_{A_1}, m'_{A_1}—all O's of A_1
m''_{A_1}
Account A_2
$\Theta_{A_2}=\{a_2,\sim a_2\}$
&
Audit Objectives for A_2
m_{A_2}, m'_{A_2}—all O's of A_2
m''_{A_2}
Audit Objective A_3O_1
$\Theta_{A_3O_1}=\{a_3o_1,\sim a_3o_1\}$
$m_{A_3O_1}$
Financial Statement
$\Theta_F=\{f,\sim f\}$
&
Account A_3
$\Theta_{A_3}=\{a_3,\sim a_3\}$
&
Audit Objective A_3O_2
$\Theta_{A_3O_2}=\{a_3o_2,\sim a_3o_2\}$
m_F, m'_F—all A's
m'_F
m_{A_3}, m'_{A_3}—all O's of A_3
m''_{A_3}
$m_{A_3O_2}$
Audit Objective A_3O_3
$\Theta_{A_3O_3}=\{a_3o_3,\sim a_3o_3\}$
$m_{A_3O_3}$
Account A_4
$\Theta_{A_4}=\{a_4,\sim a_4\}$
&
Audit Objectives of A_4
m_{A_4}, m'_{A_4}—all O's of A_4
m''_{A_4}

The next step is to combine the above **m**-values with $\mathbf{m}_{A_3}$ at the account, as defined in general terms in equations (11) through (13), before propagating it to *F*. Since there is no conflict in the two sets of **m**-values at the account level and there also is no support for $\sim a_3$, the combined values when using Dempster's rule becomes:

$$\mathbf{m}''_{A_3}(a_3)=1-\mathbf{m}''_{A_3}(\{a_3,\sim a_3\})=1-IR_{A_3}APR_{A_3}\left[1-\prod_{i=1}^{3}(1-AR_{A_3O_i})\right],$$

$$\mathbf{m}''_{A_3}(\sim a_3)=0,$$

$$\mathbf{m}''_{A_3}(\{a_3,\sim a_3\})=\mathbf{m}_{A_3}(\{a_3,\sim a_3\})\mathbf{m}'_{A_3-\text{all } O\text{'s of } A_3}(\{a_3,\sim a_3\})$$

$$=IR_{A_3}APR_{A_3}\left[1-\prod_{i=1}^{3}(1-AR_{A_3O_i})\right]. \tag{B-2}$$

In general, one can write the above result for any account '*A*' in an "and" tree similar to that in figure 1 as:

$$\mathbf{m}''_A(a)=1-\mathbf{m}''_A(\{a,\sim a\})=1-IR_AAPR_A\left[1-\prod_{O}(1-AR_{AO})\right],$$

$$\mathbf{m}''_A(\sim a)=0,$$

$$\mathbf{m}''_A(\{a,\sim a\})=IR_AAPR_A\left[1-\prod_{O}(1-AR_{AO})\right]. \tag{B-3}$$

Let us define a new term, AR_A, for simplicity as:

$$AR_A=IR_AAPR_A\left[1-\prod_{O}(1-AR_{AO})\right]. \tag{B-4}$$

To combine all the items of evidence at F, we propagate the *m*-values in equation (B-3) to *F* (see fig. 2). Using equations (A-1), (A-3), and (A-4), we can write these values, in general form, as:

$$\mathbf{m}'_{F-\text{all } A\text{'s}}(f)=\prod_{A}\mathbf{m}''_A(a)=\prod_{A}(1-AR_A),$$

$$\mathbf{m}'_{F-\text{all } A\text{'s}}(\sim f)=0,$$

$$\mathbf{m}'_{F-\text{all } A\text{'s}}(\{f,\sim f\})=1-\prod_{A}(1-AR_A). \tag{B-5}$$

Now, to complete the process, the above **m**-values are combined with $\mathbf{m}_F$ at F as defined in equations (8) through (10). Using Dempster's rule to combine $\mathbf{m}'_{F-\text{all } A\text{'s}}$ in equation (B-5) and $\mathbf{m}_F$, we obtain the total **m**-values as:

$$\mathbf{m}^t_F(f)=1-\mathbf{m}^t_F(\{f,\sim f\})=1-IR_FAPR_F\left[1-\prod_{A}(1-AR_A)\right],$$

$$\mathbf{m}^t_F(\sim f)=0,$$

$$\mathbf{m}^t_F(\{f,\sim f\})=\mathbf{m}_F(\{f,\sim f\})\mathbf{m}'_{F-\text{all } A\text{'s}}(\{f,\sim f\})=IR_FAPR_F\left[1-\prod_{A}(1-AR_A)\right]. \tag{B-6}$$

This yields the following values for beliefs and plausibilities that are of interest (using eqs. [4], [5], and [B-6]):

$$Bel^t_F(f)=1-IR_FAPR_F\left[1-\prod_{A}(1-AR_A)\right],$$

$$Bel^t_F(\sim f)=0,$$

$$PL^t_F(\sim f)=AR^t_F=IR_FAPR_F\left[1-\prod_{A}(1-AR_A)\right]. \tag{B-7}$$

Since we termed plausibility of error as audit risk, the desired formula for the audit risk is given by AR^t_F in equation (B-7), where:

$$AR_A = IR_A APR_A \left[1 - \prod_O (1 - AR_{AO})\right]$$

as defined in equation (B-4) and $AR_{AO} = IR_{AO} APR_{AO} CR_{AO} DR_{AO}$ as defined in equation (16).

Accounts Level

In this section, we derive a formula for the total plausibility of material misstatement, or the total audit risk, at the account level. Similar to the procedure for deriving the formula for the financial statement level, this derivation combines the eight items of evidence: four at the audit objective level, two at the account level, and two at the financial statement level. This will be achieved by propagating **m**-values from F, from the audit objectives, and from other accounts to the desired account. We will use the relationships developed in appendix A.

Let us consider, for the purpose of illustration, account A_3 in figure 3. The **m**-values received by A_3 from the audit objectives were determined in the previous section (eq. [B-1]). The **m**-values received from the financial statement node is given below (use eqs. [8]–[10], [A-7], [A-9], and [A-10]):

$$\begin{aligned} \mathbf{m}'_{A_3,-F\,\&\,\text{all other }A\text{'s}}(a_3) &= \mathbf{m}_F(f) = 1 - IR_F APR_F,\\ \mathbf{m}'_{A_3,-F\,\&\,\text{all other }A\text{'s}}(\sim a_3) &= 0,\\ \mathbf{m}'_{A_3,-F\,\&\,\text{all other }A\text{'s}}(\{a_3, \sim a_3\}) &= 1 - \mathbf{m}_F(f) = IR_F APR_F. \end{aligned} \tag{B-8}$$

Similar **m**-values will be received by the other accounts. At A_3, we now have three **m**-values: $\mathbf{m}_{A_3}$ defined at A_3 (eqs. [11]–[13]), $\mathbf{m}'_{A_3,-F\,\&\,\text{all other }A\text{'s}}$ as given in equation (B-8), and $\mathbf{m}'_{A_3,-\text{all }O\text{'s of }A_3}$ in equation (B-1). We combine these **m**-values using Dempster's rule, which yields the following total **m**-values:

$$\begin{aligned} \mathbf{m}'_{A_3}(a_3) &= 1 - IR_F APR_F IR_{A_3} APR_{A_3}\left[1 - \prod_{i=1}^{3}(1 - AR_{A_3O_i})\right],\\ \mathbf{m}'_{A_3}(\sim a_3) &= 0,\\ \mathbf{m}'_{A_3}(\{a_3, \sim a_3\}) &= IR_F APR_F IR_{A_3} APR_{A_3}\left[1 - \prod_{i=1}^{3}(1 - AR_{A_3O_i})\right]. \end{aligned} \tag{B-9}$$

In general, one can write equation (B-9) as:

$$\begin{aligned} \mathbf{m}'_A(a) &= 1 - IR_F APR_F IR_A APR_A\left[1 - \prod_O (1 - AR_{AO})\right],\\ \mathbf{m}'_A(\sim a) &= 0,\\ \mathbf{m}'_A(\{a, \sim a\}) &= IR_F APR_F IR_A APR_A\left[1 - \prod_O (1 - AR_{AO})\right]. \end{aligned} \tag{B-10}$$

These values yield the following beliefs and plausibility values (using eqs. [4], [5], and [B-10]):

$$\begin{aligned} Bel'_A(a) &= 1 - IR_F APR_F IR_A APR_A\left[1 - \prod_O (1 - AR_{AO})\right],\\ Bel'_A(\sim a) &= 0,\\ Pl'_A(\sim a) &= AR'_A = IR_F APR_F IR_A APR_A\left[1 - \prod_O (1 - AR_{AO})\right]. \end{aligned} \tag{B-11}$$

Since the total plausibility in $\sim a$ represents the total audit risk, AR'_A in equation (B-11) is the desired formula at the account level.

Audit Objective Level

In this section, we want to combine, again, all the items of evidence in a network in figure 1, but at the audit objective level. First, we propagate **m**-values from the financial statement node to the accounts. As an illustration, we consider first propagation of $\mathbf{m}_F$ to A_3 (see fig. 4). This yields $\mathbf{m}'_{A_3,-F\,\&\,\text{all other }A\text{'s}}$ given in equation (B-8). As a second step, we combine $\mathbf{m}'_{A_3,-F\,\&\,\text{all other }A\text{'s}}$ with $\mathbf{m}_3$ at A_3 (eqs. [11]–[13]). The result of combination yields $\mathbf{m}''_{A_3}$:

Figure 3

Propagation of m-Values from the Financial Statement and Audit Objectives to the Accounts

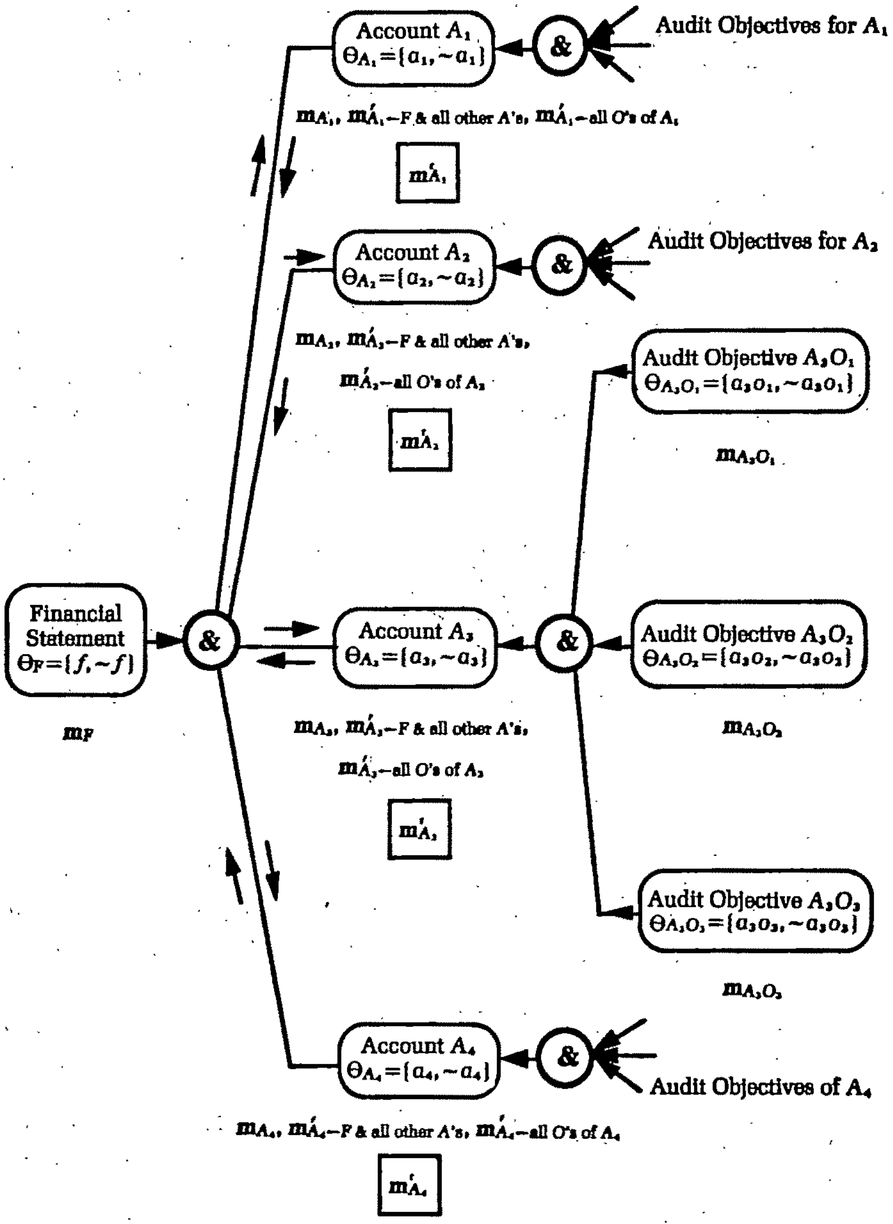
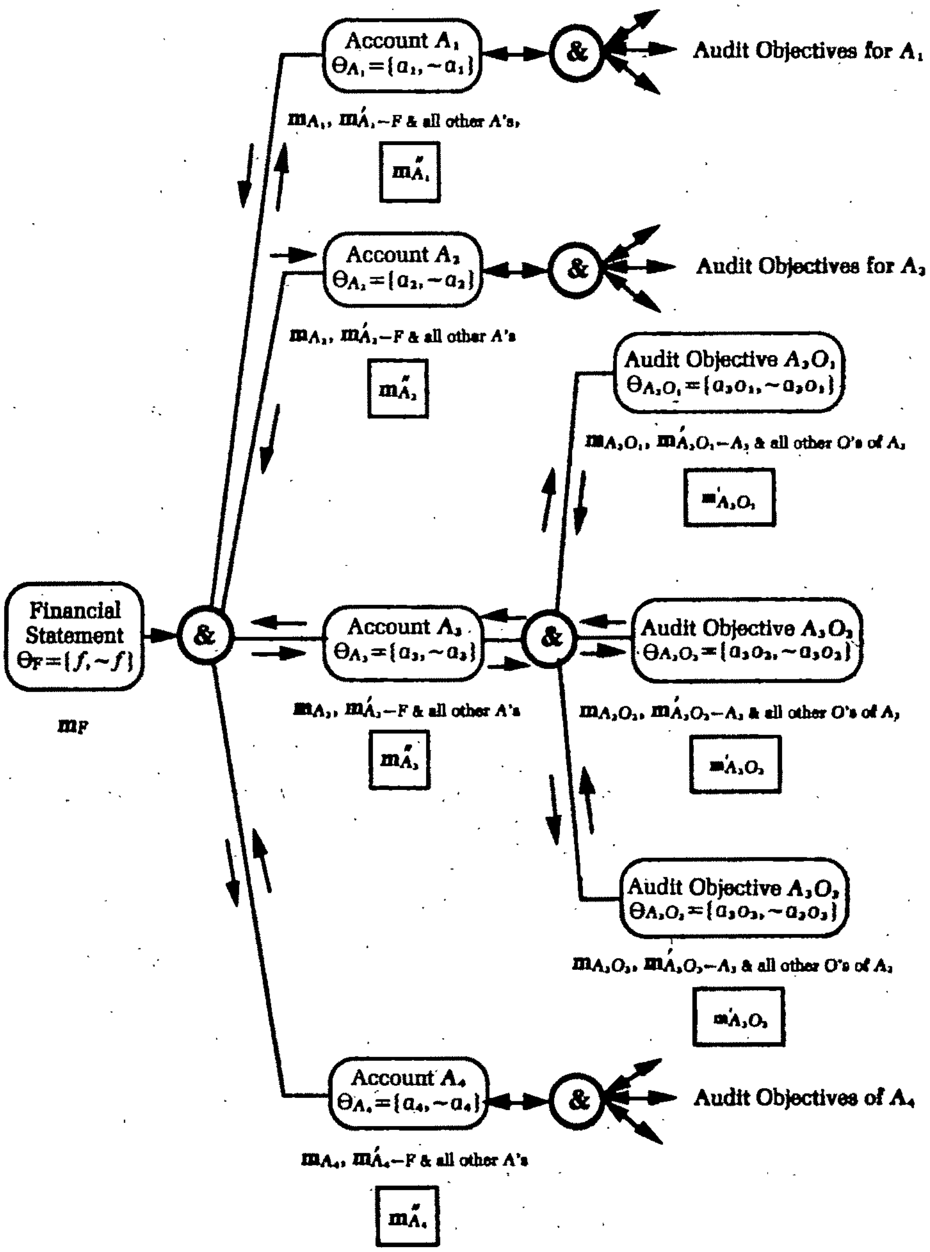

Account A_1 $\Theta_{A_1}=\{a_1,\sim a_1\}$
&
Audit Objectives for A_1
m_{A_1}, m'_{A_1}–F & all other A's, m'_{A_1}–all O's of A_1
m'_{A_1}
Account A_2 $\Theta_{A_2}=\{a_2,\sim a_2\}$
Audit Objectives for A_2
m_{A_2}, m'_{A_2}–F & all other A's,
m'_{A_2}–all O's of A_2
m'_{A_2}
Audit Objective A_3O_1 $\Theta_{A_3O_1}=\{a_3o_1,\sim a_3o_1\}$
$m_{A_3O_1}$
Financial Statement $\Theta_F=\{f,\sim f\}$
m_F
Account A_3 $\Theta_{A_3}=\{a_3,\sim a_3\}$
Audit Objective A_3O_2 $\Theta_{A_3O_2}=\{a_3o_2,\sim a_3o_2\}$
m_{A_3}, m'_{A_3}–F & all other A's,
m'_{A_3}–all O's of A_3
$m_{A_3O_2}$
m'_{A_3}
Audit Objective A_3O_3 $\Theta_{A_3O_3}=\{a_3o_3,\sim a_3o_3\}$
$m_{A_3O_3}$
Account A_4 $\Theta_{A_4}=\{a_4,\sim a_4\}$
Audit Objectives of A_4
m_{A_4}, m'_{A_4}–F & all other A's, m'_{A_4}–all O's of A_4
m'_{A_4}

Figure 4
Propagation of m-Values from the Financial Statement and Accounts to the Audit Objectives


Account A_1 $\Theta_{A_1}=\{a_1,\sim a_1\}$
&
Audit Objectives for A_1
m_{A_1}, m'_{A_1}–F & all other A's,
m''_{A_1}
Account A_2 $\Theta_{A_2}=\{a_2,\sim a_2\}$
&
Audit Objectives for A_2
m_{A_2}, m'_{A_2}–F & all other A's
m''_{A_2}
Audit Objective A_3O_1 $\Theta_{A_3O_1}=\{a_3o_1,\sim a_3o_1\}$
$m_{A_3O_1}$, $m'_{A_3O_1}$–A_3 & all other O's of A_3
$m'_{A_3O_1}$
Financial Statement $\Theta_F=\{f,\sim f\}$
m_F
&
Account A_3 $\Theta_{A_3}=\{a_3,\sim a_3\}$
&
Audit Objective A_3O_2 $\Theta_{A_3O_2}=\{a_3o_2,\sim a_3o_2\}$
m_{A_3}, m'_{A_3}–F & all other A's
m''_{A_3}
$m_{A_3O_2}$, $m'_{A_3O_2}$–A_3 & all other O's of A_3
$m'_{A_3O_2}$
Audit Objective A_3O_3 $\Theta_{A_3O_3}=\{a_3o_3,\sim a_3o_3\}$
$m_{A_3O_3}$, $m'_{A_3O_3}$–A_3 & all other O's of A_3
$m'_{A_3O_3}$
Account A_4 $\Theta_{A_4}=\{a_4,\sim a_4\}$
&
Audit Objectives of A_4
m_{A_4}, m'_{A_4}–F & all other A's
m''_{A_4}

$$\mathbf{m}''_{A_3}(a)=1-IR_FAPR_FIR_{A_3}APR_{A_3},$$

$$\mathbf{m}''_{A_3}(\sim a)=0,$$

$$\mathbf{m}''_{A_3}(\{a,\sim a\})=IR_FAPR_FIR_{A_3}APR_{A_3}. \qquad \text{(B-12)}$$

These $\mathbf{m}$-values are propagated to the various audit objectives. We use equations (A-7), (A-9), and (A-10) for this purpose. For audit objective A_3O_2, the result of propagation yields $\mathbf{m}'_{A_3O_2-A_3\ \&\ \text{all other } O\text{'s of } A_3}$:

$$\mathbf{m}'_{A_3O_2-A_3\ \&\ \text{all other } O\text{'s of } A_3}(a_3o_2)=\mathbf{m}''_{A_3}(a_3)=1-IR_FAPR_FIR_{A_3}APR_{A_3},$$

$$\mathbf{m}'_{A_3O_2-A_3\ \&\ \text{all other } O\text{'s of } A_3}(\sim a_3o_2)=0,$$

$$\mathbf{m}'_{A_3O_2-A_3\ \&\ \text{all other } O\text{'s of } A_3}(\{a_3o_2\sim a_3o_2\})=IR_FAPR_FIR_{A_3}APR_{A_3}. \qquad \text{(B-13)}$$

As shown in figure 4, we have two $\mathbf{m}$-values at each audit objective. For objective O_2 of account A_3, we have $\mathbf{m}'_{A_3O_2-A_3\ \&\ \text{all other } O\text{'s of } A_3}$ in equation (B-13) and $\mathbf{m}_{A_3O_2}$ in equations (14) through (17). We combine these values using Dempster's rule, which yields the total $\mathbf{m}$-values at the audit objective level. This represents an aggregation of all the items of evidence bearing on various levels.

$$\mathbf{m}'_{A_3O_2}(a_3o_2)=1-IR_FAPR_FIR_{A_3}APR_{A_3}AR_{A_3O_2},$$

$$\mathbf{m}'_{A_3O_2}(\sim a_3o_2)=0,$$

$$\mathbf{m}'_{A_3O_2}(\{a_3o_2,\sim a_3o_2\})=IR_FAPR_FIR_{A_3}APR_{A_3}AR_{A_3O_2}. \qquad \text{(B-14)}$$

From equations (4) and (5), one obtains the following values for beliefs and plausibilities for any audit objective AO:

$$\mathbf{Bel}'_{AO}(ao)=1-IR_FAPR_FIR_AAPR_AAR_{AO}, \qquad \text{(B-15)}$$

$$\mathbf{Bel}'_{AO}(\sim ao)=0, \qquad \text{(B-16)}$$

$$\mathbf{PL}'_{AO}(\sim ao)=AR'_{AO}=IR_FAPR_FIR_AAPR_AAR_{AO}. \qquad \text{(B-17)}$$

Equation (B-17) is the desired formula for this case.

References

American Institute of Certified Public Accountants. 1983. *Statement on Auditing Standards No. 47: Audit Risk and Materiality in Conducting an Audit.* New York: AICPA.

———. 1988a. *Statement on Auditing Standards, No. 55: Consideration of the Internal Control Structure in a Financial Statement Audit.* New York: AICPA.

———. 1988b. *Statement on Auditing Standards, No. 56: Analytical Procedures.* New York: AICPA.

Akresh, A. D., J. K. Loebbecke, and W. R. Scott. 1988. Audit approaches and techniques. In *Research Opportunities in Auditing: The Second Decade,* edited by A. R. Abdel-khalik and Ira Solomon. Sarasota, FL: AAA, 13–55.

Arens, A. A., and J. K. Loebbecke. 1988. *Auditing: An Integrated Approach.* Englewood Cliffs, NJ: Prentice-Hall.

Boritz, J. E. 1990. Appropriate and inappropriate approaches to combining evidence in an assertion-based auditing framework. Working paper, School of Accountancy, University of Waterloo, Canada.

———, and R. E. Jensen. 1985. An hierarchical, assertion-oriented approach to planning audit evidence-gathering procedures. Presented at the *Symposium on Audit Judgment and Evidence Evaluation,* University of Southern California.

———, and A. K. P. Wensley. 1990. Structuring the assessment of audit evidence: An expert system approach. *Auditing: A Journal of Practice & Theory* 9 (Supplement): 49–109.

Buchanan, B. G., and E. H. Shortliffe. 1984. *Rule-Based Expert Systems: The MYCIN Experiments of the Stanford Heuristic Programming Project.* Reading, MA: Addison-Wesley.

Canadian Institute of Chartered Accountants. 1980. *Extent of Audit Testing.* Toronto, Canada: CICA.

Cushing, B. E., and J. K. Loebbecke. 1983. Analytical approaches to audit risk: A survey and analysis. *Auditing: A Journal of Practice & Theory* 3 (Fall): 23–41.

Graham, L. E. 1985a. *Audit risk—Part I. The CPA Journal* 55 (August): 12–21.

——. 1985b. Audit risk—Part II. *The CPA Journal* 55 (September): 34–41.

——. 1985c. Audit risk—Part III. *The CPA Journal* 55 (October): 36–43.

——. 1985d. Audit risk—Part IV. *The CPA Journal* 55 (November): 38–47.

——. 1985e. Audit risk—Part V. *The CPA Journal* 55 (December): 26–35.

Kinney, Jr., W. R. 1984. Discussant's response to an analysis of the audit framework focusing on inherent risk and the role of statistical sampling in compliance testing. *Proceedings of the 1984 Touche Ross/University of Kansas Symposium on Auditing Problems.* Lawrence, KS: School of Business, University of Kansas, 126–32.

——. 1989. Achieved audit risk and the audit outcome space. *Auditing: A Journal of Practice & Theory* 8 (Supplement): 67–97.

Leslie, D. A. 1984. An analysis of the audit framework focusing on inherent risk and the role of statistical sampling in compliance testing. *Proceedings of the 1984 Touche Ross/University of Kansas Symposium on Auditing Problems.* Lawrence, KS: School of Business, University of Kansas, 89–125.

——, S. J. Aldersley, D. J. Cockburn, and C. J. Reiter. 1986. An assertion-based approach to auditing (discussant's remarks). *Proceedings of the 1986 Touche Ross/University of Kansas Symposium on Auditing Problems.* Lawrence, KS: School of Business, University of Kansas, 31–67.

Shafer, G. 1976. *A Mathematical Theory of Evidence.* Princeton University Press.

——. 1982. Belief functions and parametric models. *Journal of the Royal Statistical Society* 44 (Series B): 322–52.

——. 1986. Probability judgment in artificial intelligence and expert systems. *Statistical Science* 2 (1): 3–16.

——, and P. P. Shenoy. 1988. Local computation in hypertrees. Working Paper No. 201, School of Business, University of Kansas, Lawrence, KS.

——, and ——. 1990. Probability propagation. *Annals of Mathematics and Artificial Intelligence* 2 (1–4): 327–52.

——, ——, and R. P. Srivastava. 1988. Auditor's assistant: A knowledge engineering tool for audit decisions. *Proceedings of the 1988 Touche Ross/University of Kansas Symposium on Auditing Problems.* Lawrence, KS: School of Business, University of Kansas, 61–84.

——, and R. P. Srivastava. 1990. The Bayesian and belief-function formalisms: A general perspective for auditing. *Auditing: A Journal of Practice & Theory* 9 (Supplement): 110–48.

Shenoy, P. P. 1989. A valuation-based language for expert systems. *International Journal of Approximate Reasoning* 3 (5): 383–411.

——, and G. Shafer. 1988. An axiomatic framework for Bayesian and belief-function propagation. *Proceedings of the Fourth Workshop on Uncertainty in Artificial Intelligence.* St. Paul, MN: AAAI Press, 307–14.

——, and ——. 1990. Axioms for probability and belief-function propagation. In *Uncertainty in Artificial Intelligence 4*, edited by R. D. Shachter et al. Amsterdam: North-Holland, 169–98.

Zarley, D., Yen-Teh Hsia, and G. Shafer. 1988. Evidential reasoning using DELIEF. *Proceedings of the Seventh National Conference on Artificial Intelligence*, Vol. 1. Cambridge, MA: AAAI Press, 205–09.

THE ACCOUNTING REVIEW
Vol. 67, No. 2
April 1992
pp. 284–302

A Multidimensional Analysis of Selected Ethical Issues in Accounting

Steven M. Flory
Thomas J. Phillips, Jr.
Louisiana Tech University
R. Eric Reidenbach
Donald P. Robin
University of Southern Mississippi

SYNOPSIS AND INTRODUCTION: Much of the past research in accounting ethics has focused on whether accountants conform to prescribed codes of professional ethics. Other research has been normative in nature, recommending what constitutes appropriate ethical conduct or focusing on the accountant's responsibility in society. This study selects a different approach by testing a multivariate measure of how accountants make ethical judgments.

Data were gathered with the assistance of the Institute of Certified Management Accountants and the Institute of Management Accountants (formerly the National Association of Accountants). Accountants were asked to respond on bipolar scales to realistic scenarios involving ethical decisions. Several tests for construct validity produced supportive results for the hypothesized three-dimensional measure, with the dimensions being moral equity, relativism, and contractualism.

First, we developed a questionnaire with four scenarios concerning ethical issues. Each scenario ended in a particular action taken by an individual. Responses to that action were recorded on eight bipolar scales, representing the three dimensions above. The questionnaire was mailed to 500 randomly selected certified management accountants resulting in a 62.8 percent response rate.

The authors wish to thank the anonymous referees for their helpful comments. We are indebted to Gary Scopes, executive director and Alfred King, both of the Institute of Management Accountants and James Bullock, director of the Institute of Certified Management Accountants for their valuable cooperation in the data gathering stage. In addition, we thank Awni Zebda and Otis Gilley for their useful suggestions. Each of the four authors made equal contributions.

Submitted January 1991.
Accepted September 1991.

Second, the results from a factor analysis and a traditional reliability coefficient test suggest that a high degree of internal consistency exists for each dimension of the measure. The appropriate factor loadings ranged from a low of 0.68 to a high of 0.92, while the reliability coefficients varied from 0.75 to 0.92.

Next, the content validity of the three-dimensional measure was checked by comparing it with a global ethical/unethical measure. Again, the results support the hypothesis that the multivariate measure captures the appropriate domain of content. Adjusted R^2-values ranged from 0.59 to 0.76 when the global measure was regressed against the multivariate measure.

Finally, a sense of predictive validity was obtained by comparing the multivariate measure with a behavioral intention measure for the respondent. Adjusted R^2-values ranged from 0.45 to 0.76 for the four scenarios tested, indicating that the three-dimensional measure "explains" a respectable portion of the variance in the behavioral intention of the individual.

The multidimensional measure developed in this study may be a guide for future research into how accountants make ethical judgments. Such knowledge can be used in turn to develop useful codes of conduct, create ethical organizational cultures, and direct ethical training for and by accountants.

Key Words: *Accounting ethics research, Multidimensional ethics model, Psychometric analysis, Behavioral ethics measures.*

Data Availability: *The data upon which this article is based may be obtained from the authors on request.*

THE remainder of this paper consists of the following. Section I reviews the literature on the development of codes of ethics and examines research regarding the role of the accounting profession's responsibility to society. Section II is concerned with the development of the multidimensional ethics measure used in the research. The survey is described in section III, followed by the results in section IV that consider the internal consistency of the constructs, content validity, and predictive validity of the ethics measure. Section V provides a comparison of the scenarios. The final section contains our conclusions.

I. Background

Much of the literature in accounting ethics is prescriptive regarding a satisfactory code of ethics for practitioners. However, some recent studies are empirically-based, emphasizing the context of accounting ethics. These studies focus heavily on legitimatizing the accounting profession's status in society, some considering the accountant's responsibilities and others proposing needed social changes. While this research concentrates instead on the dynamics of the ethical decision-making process, it draws from some of the concepts found in these recent studies. This article develops along the lines of a positive theory of ethical decision-making. The background begins with an examination of the prescriptive literature and proceeds through the recent, empirically-based studies.

Social concern regarding accounting ethics in the United States, United Kingdom, and Australia has developed along similar lines by focusing on the interaction of ethics and professionalism and by emphasizing the importance of self-regulation (Carey 1946; Parker 1986; Parker 1987). Most authors agree (e.g., Loeb 1988; Mautz 1984, 1988) that self-regulation is in the best interest of the profession. Mautz argues that the profession is given a special franchise in return for accepting responsibility to society.

Throughout the development of codes of ethics, from the initial American Institute of Accountants' code in 1917 to the recently adopted American Institute of Certified Public Accountants (AICPA) Code of 1988, practitioners have stressed four areas of interest to the profession and public. Of prime importance to the structuring of codes are: "... (1) independence, (2) advertising and solicitation, (3) commissions, and (4) contingent fees" (Lowe 1987, 80). Because all these issues can negatively affect public confidence in the profession, they are strongly interrelated with the profession's self-interest.

Many of the early discussions of accounting ethics focused on the need to establish public confidence in the accounting profession through ethics codes such as those governing practice in the legal and medical professions (Carey 1946, 1947; Richardson 1931). This image-conscious, practical approach to specific ethical issues as they relate to current accounting practice is prevalent throughout the professional literature of both the United States and other countries (Parker 1987; Woelfel 1986).

A major difference between the influence of ethics in the legal and medical professions, on the one hand, and accountants, on the other hand, is that the first obligation of lawyers and doctors is to the client, while professional accountants must satisfy a broader obligation to third parties who rely on their reports. Despite this distinction, ethics research in accounting has generally centered around the establishment of pragmatic, rules-oriented codes intended to regulate public accounting practice, largely avoiding philosophical discussions about good and evil because "... practical men feel that they have no time for these luxuries" (Carey 1947, 122).

Practicing CPAs asked to describe situations that posed the most difficult ethical situations focused on specific practice issues, such as client pressure to alter tax reports, conflict of interest/independence situations, and client fee problems (Finn et al. 1988). Respondents also indicated that, although there were numerous opportunities for CPAs to engage in unethical behavior, few of them did so frequently, and that the AICPA Code of Ethics effectively addressed the ethical issues of greatest concern to practicing CPAs.

The effect of self-interest on the formulation and enforcement of ethical standards has been an area of consistently strong criticism in accounting literature (Briloff 1972, 1976, 1981, 1986a, 1986b; Davis 1984; Graber 1979; Lehman 1988; Tinker 1988). Loeb (1971, 1972) studied sanctions imposed for ethics violations and conducted a survey of ethical behavior among CPAs. In an examination of ethics violation charges heard by the Wisconsin State Board of Public Accountancy and the State Society of CPAs, Loeb found that offenses against colleagues were the predominant violation cited and that ethics violations toward clients were seldom disciplined. A comparison with ethics code sanctions in the legal profession revealed similarities in the pattern of penalties imposed.

A second part of the Loeb research, testing acceptance of ethical norms in various ethical conflict situations, led to the conclusion that "... in a profession relying on self policy, adherence to ... a code of ethics depends to a substantial degree on the broad acceptence of the provisions of the code" (Loeb 1971, 301). Similar opinions were expressed in a survey of business executives, academic deans, and members of Congress (Touche Ross 1988). Respondents to this survey supported self-regulation through adoption of business codes of ethics as the most effective measure for encouraging ethical business behavior; they named government legislation as the least effective means of doing so. Both of these responses lend support to the traditional view of ethics as a self-directed instrument capable of facilitating normal business practice. In addition, approximately two-thirds of the respondents expressed an opinion that "... high ethical standards in a business enterprise strengthen its competitive position" (Touche Ross 1988, 3).

Public perception of the accounting profession's ability to regulate itself has sometimes been negative, as evidenced by criticism from congressional leaders such as Ron Wyden and John Dingell. The Treadway Commission Report (National Commission on Fraudulent Financial Reporting 1987) was an attempt to improve the profession's response to the public interest, especially in regard to decreasing the incidence of fraudulent reporting. Treadway recognized a need for integrating ethical values into accounting and business coursework, as did the Bedford Report (American Accounting Association 1986).

Treadway also emphasized management's responsibility in financial reporting for setting the proper "tone at the top" for ethical conduct, stating that a company should establish a code of conduct for its employees to follow. Magill and Previts (1991), however, warned against establishing codes without recognizing the need for a commitment to serve others because such rules may come to be viewed as reflecting acceptable behavior instead of setting minimum ideals.

A number of researchers have expanded their view of accounting thought to include the broad relationships and influences of social, political, and economic factors upon the role of accounting (Burchell et al. 1980; Colignon and Covaleski 1991; Gaa 1986; Gray 1988; Lev 1988; Noreen 1988; Tinker 1988; Tinker et al. 1982). They advocate awareness of the pervasiveness of these influences on accounting rather than a search for an elusive, perhaps nonexistent, fundamental truth.[1]

This viewpoint is drawn from positive accounting theory (Watts and Zimmerman 1978, 1986), which recognizes the "self-interest" motive in financial reporting as opposed to the reporting of unbiased economic reality. Gaa (1986, 435), for instance, states that "... [t]he collective choice aspects of financial reporting are the unavoidable result of the existence of conflicting individual interests." He argues that in the absence of a "best" choice, reporting is often biased toward the public interest, although he questions whether this bias is soundly based.

A public interest bias, however, is consistent with Lehman's (1988) view that there should be a redefinition of accounting ethics that encompasses an extended sociopolitical role of accounting in society. In a study of articles published in three prominent business journals, Lehman found an emphasis on competitive behavior and wealth maximization, with a corresponding lack of reference to moral and ethical behavior and social responsibility. The study concluded that this "survival of the fittest" attitude in business behavior results in an "ethics for profits" rationalization apparent in the results of the Touche Ross study mentioned earlier.

---

[1] An opposing view is one that encompasses the concept of fairness (e.g., see Solomons 1991; Williams 1987) and sees accounting as an attempt to uncover a faithful, unbiased representation of economic truth. Accepting the concept that there is a single, fundamentally correct representation of economic reality does not diminish the need for our study. Our attempt here is to understand individuals' ethical behavior, regardless of what one argues that behavior should be.

Regardless of one's assessment of the desirability of the sociopolitical policies proposed by Lehman and others (Benston 1982; Lehman and Tinker 1987; Merino and Neimark 1982; Tinker 1985; Tinker et al. 1982), their concern focuses on the policies and institutional changes needed to achieve desired social goals rather than on the processes and variables involved in making ethical decisions. This needed approach to ethics, broader in scope than earlier attempts, is not directed toward an understanding of ethical behavior.

A non-prescriptive study of professional responsibilities by Gaa (1990) showed that the opportunities for professionals to act in their own self-interest and disregard their responsibilities allow structural instability in the relationship between professionals and society. In an examination of ethical behavior in the context of agency theory, Noreen (1988, 359) concluded that "... an agreement (i.e., ethical code) to abstain from opportunistic behavior cannot be enforced effectively by external rewards or sanctions; instead, the sanctions for unethical behavior must be internalized." The process of internalizing ethical sanctions may be an opportunity to stabilize the "social contract" between the accounting profession and society referred to by Gaa (1986, 1990). However, internalization of these sanctions calls for a better understanding of ethical behavior and of the multidimensional factors involved in shaping ethical conduct.

In a review of the Special Committee report (the Anderson Committee, AICPA 1986) calling for substantial changes in the AICPA Code of Ethics, Shaub (1988, 97) makes a recommendation for further research on ethical standards governing CPAs that could apply to standards affecting all professional accountants:

> Specifically, it is time for the profession to undertake a study of the determinants of ethical behavior in CPAs. This is not the same as determining whether or not CPAs are ethical. Instead, research should incorporate a body of knowledge that has been developed in the social sciences, and is currently being adapted into business disciplines such as marketing, that provides some insight into what causes individuals to *choose* ethical behavior.

Such research embraces the reasoning of positive thought in that it seeks to explain and predict ethical behavior. While a normative, prescriptive framework for ethical conduct is vital and deserves considerably more research, an adequate understanding of ethical behavior and the factors that influence behavior is imperative.

Accordingly, this study does not endeavor to delineate right from wrong. Neither does it attempt to prescribe ethical conduct. Its purpose is to investigate factors that might influence ethical behavior, realizing that these factors are potentially complex and multifaceted.

II. Development of the Multidimensional Ethics Measure

Psychometricians almost never regard a measure completely valid in every circumstance, but rather consider the process of validation to be unending. Further, measures of potentially broad constructs, such as "ethics," must be tested in a variety of settings

Figure 1
The Multidimensional Ethics Measure

Moral Equity Dimension

FAIR	_____:_____:_____:_____:_____:_____:_____	UNFAIR
JUST	_____:_____:_____:_____:_____:_____:_____	UNJUST
MORALLY RIGHT	_____:_____:_____:_____:_____:_____:_____	NOT MORALLY RIGHT
ACCEPTABLE TO MY FAMILY	_____:_____:_____:_____:_____:_____:_____	UNACCEPTABLE TO MY FAMILY

Relativism Dimension

CULTURALLY ACCEPTABLE	_____:_____:_____:_____:_____:_____:_____	CULTURALLY UNACCEPTABLE
TRADITIONALLY ACCEPTABLE	_____:_____:_____:_____:_____:_____:_____	TRADITIONALLY UNACCEPTABLE

Contractualism Dimension

VIOLATES AN UNWRITTEN CONTRACT	_____:_____:_____:_____:_____:_____:_____	DOES NOT VIOLATE AN UNWRITTEN CONTRACT
VIOLATES AN UNSPOKEN PROMISE	_____:_____:_____:_____:_____:_____:_____	DOES NOT VIOLATE AN UNSPOKEN PROMISE

and with different groups to confirm that the initial results are not an artifact of a particular group or setting.

The multidimensional measure developed by Reidenbach and Robin (1990) that appears in figure 1 is used in this study to focus on previously unstudied dynamics of decision making regarding ethical issues in accounting. The measure consists of eight bipolar scales divided into three dimensions—moral equity, relativism, and contractualism. The scenarios that Reidenbach and Robin used were short descriptions of single-issue situations, averaging 100 words per scenario. While this approach provided a useful first step in developing a measure of behavior, such simplified scenarios seem inadequate for complex, realistic ethical problems. The scenarios used in this study average 200 words and portray substantially more involved, realistic situations. Additionally, the initial tests were conducted with retail managers, who have a somewhat different set of demands operating on them from those affecting professional accountants. Validation of this measure in the accounting environment would add to its applicability.

Results of pretests conducted through a statewide CPA organization suggested that the dimensions developed by Reidenbach and Robin might be usable to analyze responses of accountants, and those preliminary results provide the impetus for the more substantial test in this study. Thus, it is suggested that accountants use the moral equity, relativism, and contractualism dimensions for evaluating ethical issues in accounting.

A potential concern in developing a multidimensional ethics measure has been addressed in previous social psychology literature. Hartshorne and May (1928), for instance, found situational factors to be more influential of ethical behavior than personal characteristics and traits. Their results from testing thousands of children failed to show morality as a consistent, general trait. This classic study in human honesty has been replicated by other social psychologists (e.g., Brown et al. 1969; Hoffman 1977) whose results support the Hartshorne and May study. Brown et al., for instance, used college students in their replication to avoid the criticism that morality may be undeveloped prior to adolescence. Fletcher (1966) also advanced the idea that ethical decisions vary by situation, although some psychologists have shown weak evidence that a general morality factor exists (Burton 1963; Nelson et al. 1969).

These findings tend to complicate efforts to understand and predict behavior. The present study attempts to delineate the factors that influence judgments about ethical behavior in different situations. It is intended as a first step to show that the same factors influence ethical behavior regardless of the individual's decisions in a particular situation.

III. The Survey

Practical scenarios using realistic ethical problems faced by accountants, as well as responses from a nationwide group of accounting practitioners, were necessary for the success of this study. Both problems were addressed by the participation of the Institute of Management Accountants (IMA, formerly the National Association of Accountants).

The IMA had contracted with the Ethics Resource Center in Washington, D.C., to produce a videotape that included five substantial ethical situations for accountants. The authors were given permission to abstract the scenarios for this study from that videotape. Four of the five ethical situations were chosen, excluding one somewhat redundant situation.

The Institute of Certified Management Accountants (ICMA, the certification body of the IMA) agreed to sponsor the study to the extent of providing a cover letter for a mail survey asking its certified members to participate. A mail survey was selected over other options because of the ego-involving character of some of the questions and the halo effects that often occur in responses to ethics questions. Providing anonymity to respondents in a mail survey tends to dampen these effects.

The abstracted scenarios averaged 200 words and retained the essential complexity of the ethical problem in the original videotape. Each scenario included an action statement to assure that all respondents were reacting to the same stimulus. The action statement was particularly necessary with the more complex situations described in this research. The scenarios used in the test instrument appear in the appendix along with a description of questionnaire construction. Each scenario portrays a different sort of ethical dilemma. Scenarios A and D describe actions that might not be perceived as explicitly ethical or unethical, while scenarios B and C feature what most would label as definitely unethical behavior.

Scenario A describes a superior who is making questionable expenditures that he claims meet upper management's approval. The accountant, who finds himself in a marketing environment different from his background, is faced with establishing the proper lines of authority in connection with an issue that may not be a violation of company policy.

Scenario B involves a controller who is asked to falsify external financial statements for the purpose of procuring additional working capital. Although this may actually happen in some companies, accountants typically agree that falsification of external statements is wrong. This is also true of the specific violations of company policy shown in scenario C. A difference in scenario C, besides the fact that it is an internal situation, is that the accountant had previously violated company policy, and now, in an attempt to rectify a resulting failure, decides to violate the policy again.

Scenarios A, B, and C all implicitly involve an accountant's job security, but in each situation the individuals are seemingly concerned with their company's welfare; in contrast, scenario D emphasizes the accountant's personal problems. In this scenario, company policy is not clearly delineated, and there could be some uncertainty whether the accountant's action is unethical. The additional background information provided in scenario D allows the respondent to empathize with the accountant's personal difficulties, although it is unclear whether his personal situation has any bearing on his decision.

Previous research (Loeb 1971) has shown demographic characteristics to be insignificant in their relation to ethical decisions. The certified management accountants (CMAs) chosen for the test population represent a broad cross-section of professional accountants employed in private industry, nonprofit organizations, and public practice. A mailing of the test instrument to a randomly selected sample of 500 CMAs produced 314 usable responses, a 62.8 percent return rate. One follow-up mailing was used to help achieve these results. In order to produce the most reliable responses to sensitive questions regarding ethics, respondent anonymity was promised and maintained.

Nunnally (1978, 87) states that there are "three major functions" of measures such as the one proposed here and that these functions "correspond" to three types of tests for validity. They are: (1) the measure of psychological traits or dimensions—construct validity, (2) a representation of the complete universe or domain of the construct—content validity, and (3) the establishment of a theoretically justified statistical relationship with another variable—predictive validity. The three sections that follow describe relevant findings for each of these types of validity.

IV. Results

The ethical evaluation of a situation is an abstract concept and cannot be measured directly (e.g., with a ruler or a balance). Psychometricians label such concepts "constructs," and they have developed indirect methods for measuring them.

Constructs vary widely in the number of observables that can be used to represent them; those with a considerable variety, depth, and number of potential measures are said to have a high degree of complexity. Repeated use of the same measurement items by different individuals to evaluate different situations or applications is an important test of the construct's validity (Nunnally 1978, 94–105).

Ethical judgment is a construct that is likely to have considerable complexity. How does an individual make an ethical judgment? What are the components or dimensions of such a judgment? Are the same decision criteria always used?

For the purposes of this research, it is hypothesized that accountants are influenced by the three dimensions shown in figure 1: moral equity, relativism, and contractualism. Each dimension represents a part of the construct that accountants use in making an ethical judgment, and each is measured by more than one scale in an attempt to capture its complexity.

Internal Consistency of the Scales

One approach to determining construct validity focuses on the reliability of the scales. If the dimensions operate as hypothesized, the eight scale items used to represent them should group together naturally when analyzed. Specifically, a factor analysis of the eight scale items should produce three factors, each with the anticipated scales for that dimension (see fig. 1). With four separate ethical situations used in this study, four independent factor analyses can be conducted, each expected to produce the same results. Further, the scales in each dimension should exhibit a high degree of internal consistency or "reliability."

Another test of internal consistency is factor analysis. Four factor analyses were conducted by using a principal components analysis with an orthogonal (varimax) rotation and constraining the outcome to confirm the three anticipated factors. As table 1 indicates, the anticipated dimensions (moral equity, relativism, and contractualism) are apparent in each independent run. Factor loadings ranged from the high 0.60s to the low 0.90s, and the explained variance for the four runs ranged from 77.2 to 88.1 percent. The factor loadings are distinctive enough to assuage most doubt about scale item and factor associations, and these assignments are identified with an underline in table 1.

The reliability coefficients for each dimension in each scenario are presented in table 2. Coefficient alpha, a test of consistency among scales used to measure each dimension and perhaps the most commonly used measure of reliability, ranged between 0.75 and 0.94, with an average of 0.86. These results compare favorably with those of other studies in the general business literature; they are, moreover, well above the standards of acceptability defined by Nunnally (1967, 226). Further, the values for the two-item relativism and contractualism dimensions seem particularly high, given that a smaller number of scale items often produces lower coefficient alphas.

Content Validity of the Scales

A particularly important question in establishing the validity of a measure addresses whether the construct is actually measuring what it purports to measure. That is, test scales might indicate good results from a factor analysis, reliability scores could be high, and a construct could even be a good predictor of ethical behavior, but the scale may not be measuring ethical perceptions.

Ideally, the measure of a complex construct samples from its entire domain of content. In this research, for example, all the dimensions that are ever used to make ethical judgments should be carefully identified and fully represented by an adequate number of items for each. Initial, development-stage research attempted to acknowledge this by sampling the moral philosophy literature for language used to describe ethical or

Table 1
Factor Structures for the Four Scenarios

	Scenario											
Dimension	A	B	C	D	A	B	C	D	A	B	C	D
Moral Equity												
Fair/unfair	.86	.90	.89	.88	.16	.16	.16	.29	.21	.15	.17	.26
Just/unjust	.87	.92	.88	.88	.13	.18	.22	.31	.19	.14	.21	.29
Morally right/ not morally right	.74	.75	.74	.75	.12	.23	.29	.30	.29	.34	.28	.40
Acceptable/ unacceptable to family	.69	.68	.75	.70	.34	.27	.34	.40	.11	.21	.12	.35
Relativism												
Traditionally acceptable/ unacceptable	.18	.24	.26	.33	.89	.89	.91	.89	.16	.18	.07	.17
Culturally acceptable/ unacceptable	.22	.24	.30	.33	.89	.90	.87	.88	.11	.14	.13	.21
Contractualism												
Violates/does not violate promise	.20	.18	.26	.28	.11	.13	.07	.17	.88	.90	.89	.88
Violates/does not violate contract	.28	.28	.16	.32	.16	.17	.12	.20	.82	.84	.91	.85

Note: The scale item and factor associations are underlined. The percentage of variance explained by the three-factor solution is as follows:

Scenario A, 77.2%;
Scenario B, 81.4%;
Scenario C, 82.7%;
Scenario D, 88.1%.

unethical behavior. The rationale for this approach is that the language of the moral philosophies, as well as some variation of the ideas, permeates people's expression and understanding.

To test the thesis that our constructs in fact measure ethical perception, we compared the three dimensions with a univariate ethics measure (ethical—unethical). The results appear in table 3. This is a common validation procedure in the social sciences, used here in our validation efforts. A high covariation (adjusted R^2s) between the univariate and multivariate measures suggests that the multivariate measure captures

Table 2
Reliability Coefficients (Coefficient Alpha) for Each Dimension and Each Scenario

	Scenario			
Dimension	A	B	C	D
Moral Equity	.86	.89	.90	.94
Relativism	.83	.88	.87	.92
Contractualism	.75	.81	.85	.86

Table 3
A Comparison of the Multivariate and Univariate Ethics Measures

	Regression Results			
Scenario	*Overall* R^2	B_1	B_2	B_3
A	0.59	0.76	0.43	0.48
B	0.66	0.81	0.42	0.51
C	0.66	0.80	0.52	0.42
D	0.76	0.87	0.64	0.63

Note: Overall R^2 is the result of a multivariate analysis and is adjusted for degrees of freedom. All beta values are summated measures (see fig. 1) and were calculated from individual bivariate analyses in order to focus on individual effects without colinear impacts. B_1 = beta value for moral equity dimension; B_2 = beta value for relativism dimension; B_3 = beta value for contractualism dimension. The dependent variable is the univariate measure. All results were statistically significant at <0.00005.

much of what the respondents mean by "ethical." The individual beta values for each of the three dimensions, and in each of the scenarios, also helped define the concept of "ethics" for the respondents.

As table 3 indicates, the three dimensions explain from 59 to 76 percent of the variance in what the accountants defined as ethical. The unexplained portion suggests that other items or factors might be operant in the ethical decision-making process. Moreover, in all four scenarios the moral equity dimension exerted the greatest influence. In scenarios A and B, the contractualism dimension exerted the next highest relative influence. Relativism exceeded contractualism in scenario C, and in scenario D the relativism and contractualism dimensions exerted about equal influence. This "switching" effect is examined in greater detail in a subsequent section.

Predictive Validity of the Scales

Predictive validity is an issue when the measure in question is intended for use as a predictor of some criterion or dependent variable. For this reason, it is sometimes

Table 4
The Relationship Between the Multivariate Ethics Measure and Behavioral Intention

	Regression Results[a]					*Multivariate Means*[c]			*Univariate Means*[d]
Scenario	*Overall* R^2	B_1	B_2	B_3	*Means for Behavioral Intention*[b]	*Moral Equity*	*Relativism*	*Contractualism*	*Ethical/ Unethical*
A	0.45	0.62	0.52	0.41	5.3	5.6	4.0	5.2	5.8
B	0.59	0.73	0.58	0.43	5.6	5.7	4.2	5.2	6.1
C	0.63	0.78	0.54	0.41	5.8	5.6	4.5	5.3	5.8
D	0.76	0.86	0.68	0.63	4.8	4.8	3.6	4.6	4.8

[a] B_1 = beta value for moral equity dimension; B_2 = beta value for relativism dimension; B_3 = beta value for contractualism dimension. All beta values were calculated from individual bivariate analyses in order to focus on individual effects without colinear impacts. The dependent variable is behavioral intention. All results were statistically significant at <0.00005.

[b] Behavior Intention was measured by using a single bipolar scale with the statement, "If you were responsible for making the decision described in the scenario, what is the probability that you would make the same decision?" Highly probable (scale value = 1) and highly improbable (scale value = 7) were the response options.

[c] All scales were coded from 1 to 7, and summated constructs were averaged to allow for comparison. For moral equity (fair, just, etc.), 1.0; for relativism (culturally and traditionally acceptable), 1.0; for contractualism (does not violate unwritten contract and spoken promise), 1.0.

[d] For the univariate measure, ethical = 1.0.

called criterion validity or criterion-referenced validity. A common test for predictive validity is the correspondence between two measures that are hypothesized to be related, which is usually represented by correlation (see Nunnally 1978, 88).

While it is possible that an individual's ethical response to an observed happening might not totally predict that person's willingness to behave in the same way, it is reasonable to expect ethical perceptions to play an important and significant role in that determination (see e.g., Ferrell and Gresham 1985; Hunt and Vitell 1986). Therefore, a measure of behavioral intention is the criterion variable in this test.

The theoretical basis for using behavioral intention as a surrogate measure of behavior is grounded in the work of Fishbein and Ajzen (1975), who have demonstrated empirically that intention can be used as an immediate precursor to behavior. Intention represents the individual subjective probability of engaging in the behavior. While the lack of an actual behavior measure is a limitation of the study, the concern that the ego involving character of the responses would produce a socially desirable response bias by respondents is somewhat moderated by the lack of end-loading of the behavioral intention means (table 4). These mean values denote that many respondents indicated some probability of engaging in the activity specified in the scenarios.

The results of the test for predictive validity appear in table 4. Because actual prediction cannot be achieved from cross-sectional data such as these, results from the common variation used to check on predictive validity appear in quotation marks for the remainder of the text to indicate this limitation. The common variation between constructs can be informative for the desired objectives, nevertheless.

Table 5
Difference in Explanatory Power of Multidimensional Scale Items Over the Univariate (Ethical/Unethical) Measure

	Behavioral Intention (Adjusted R^2)		
Scenario	*Univariate Measure*	*Multidimensional Measure*	*Difference*
A	0.38	0.45	+0.07
B	0.48	0.59	+0.11*
C	0.51	0.63	+0.12*
D	0.64	0.76	+0.12*

* Significant at $p<0.10$.

The behavioral intention of the respondent in response to each scenario was measured by the statement, "If you were responsible for making the decision described in the scenario, what is the probability that you would make the same decision?" The response opportunities on a seven-point bipolar scale range from highly probably to highly improbable. The relevant adjusted R^2s from "predicting" this measure with the multivariate ethics scale range from 0.45 for scenario A to 0.76 for scenario D. These results seem to satisfy the expectations for predictive validity.

The individual bivariate beta values included in table 4 are statistically significant at a 0.00005 level. One result of note is that all three dimensions play a useful explanatory role for all scenarios. The moral equity dimension (B_1) was the most dominant "predictor" of the three ethics dimensions for all scenarios; the contractualism dimension (B_3), the weakest. Further, the relative changes in R^2s, the size of the beta values, and the mean values for the different measures add additional insight into the ethical evaluation process of the respondents. Specifically, the multivariate mean values measure the respondents' ethical evaluations, while the beta values measure the relative importance of these dimensions in "predicting" behavioral intention.

The usefulness of the multivariate ethics measure may be corroborated by comparing it with the univariate ethics measure in terms of their ability to "explain" behavioral intention. The results of this comparison appear in table 5. In each case the multivariate ethics measure accounted for more "explained" variance than the univariate measure by 7 to 12 percentage points. If a one-tailed Z-test is used to determine statistically significant differences between correlations, none of the results is significant at the 95 percent level, but the differences for the B, C, and D scenarios are significant at a 90 percent level of confidence. Thus, the multidimensional scales do seem to increase our understanding of individual intention beyond what we can learn with a single-item measure.

V. A Comparison of Scenarios

One benefit of a multidimensional analysis is the possible formulation of other theoretical propositions. In this section, comparisons of findings are made between the individual scenarios, and hypotheses about the differences are offered. Comparative findings, as opposed to absolute values of dimensional averages, are the only meaningful analysis that can be conducted until norms are established for each measure.

Statistical significance is not a particularly useful tool in making these comparisons because the standard error of the means for the relevant measures ranged from 0.062 to 0.099. The small standard errors imply that even two-tenths of a scale point difference between means will be statistically significant at a 95 percent level of confidence. Meaningful differences would seem to require more of a scale point difference between means, for example, five-tenths.

One comparison of interest is between scenarios D and A. In scenario D the assistant controller of a company, who is in a position to make all purchase decisions, decides to accept a credit sale that is riskier than the company would likely approve. This scenario provides an interesting benchmark for comparison with the other three scenarios because it was judged: (1) least unethical, (2) least morally equitable, (3) most culturally/traditionally acceptable, and (4) least likely to violate an implied contract/promise, compared to the other three scenarios (see table 4). Further, respondents were most likely to say they would have behaved in the same way as the assistant controller, and the three ethical dimensions "explained" more than three-fourths of the behavorial intention score. The latter finding indicates that the more lenient evaluation of the moral equity, relativism, and implied contract dimension played an important role in how the respondents said they would behave.

In scenario A, a young management accountant approved an expense that he knows should not be approved because of pressure from the divisional vice president and lack of support from his boss. In comparison with scenario D, the respondents judged the young accountant's action as: (1) substantially more unethical, (2) less morally equitable, (3) similar in cultural acceptance, and (4) more likely to violate an implied contract. The respondents were also less likely to have behaved in the same way as the young accountant than as the assistant controller in scenario D, but were more likely to have behaved as the young accountant did than as the accounting professionals in the other two situations (B and C).

There seemed to be some sympathy with the plight of the young accountant in scenario A. This hypothesis is supported further by comparing results on the role that ethical judgment played in making the behavioral intention decision. The three ethical dimensions were least important in "explaining" behaviorial intention in scenario A than in any of the other three scenarios (adjusted $R^2=0.45$). An implication is that something else besides the ethics of the situation played an important role in influencing behavioral intention. One hypothesis is that the pressure of a situation such as this one can move people to do things that they know are unethical.

Another interesting comparison is between scenarios D and C. In scenario C, the chief accountant of the plant violated company policy to allow the purchase of a new computer system that did not perform as expected. He then agreed to spend additional funds on a system upgrade in an attempt to make the system work. In this scenario, a breach of company policy has produced a failed result (something that could occur in scenario D but has not thus far). Except for the failure of the initial action in scenario C, the two situations can be judged as similar, and thus a comparison of results allows us to develop a hypothesis about a situation that has resulted in failure, but where a similar decision might achieve the intended results and conceal the mistake.

The ethical comparison of the two scenarios produces some radical differences. The actions of the chief accountant in scenario C were judged to be: (1) more unethical, (2) less morally equitable, (3) least culturally acceptable, and (4) most likely to violate an implied contract. Respondents were least likely to undertake the same action as the chief accountant in this scenario, and the three ethical dimensions "explained" almost two-thirds of the variance of the behavioral intention measure.

Thus, respondents seemed to be saying that the action of the chief accountant in scenario C was less ethically acceptable than the assistant controller's action in scenario D, and this difference in ethical judgment seemed to be the primary reason why they were less likely to behave like the chief accountant in scenario C. One hypothesis that can be derived from these results is that failure (and potential punishment?) following an action affects the individual's ethical judgment of subsequent action.

The final comparison matches the actions of the company controller in scenario B with those in the other three scenarios. Scenario B is unique in that it is the only situation in which the accountant is asked to mislead an external public. In this scenario the controller decides to manipulate the company's books so that it appears that the company can meet its earnings forecast and raise working capital. The other three scenarios present within-company issues—the subsequent action affects only the company.

Compared to the accountants' actions in the other three scenarios, the action in scenario B is judged to be: (1) most highly unethical, (2) in approximately a three-way tie for the least morally equitable, (3) about moderate in cultural acceptability, and (4) in approximately a three-way tie for the most likely to breach an implied contract. This was the second most unlikely situation in which respondents said they would be willing to participate. The three ethical dimensions "explained" 59 percent of the variance in the behavioral intention measure. One hypothesis that could be drawn from this finding is that ethical judgment is harsher when an accountant acts in a way that can harm external publics.

VI. Conclusions

The research presented in this article tested a multidimensional measure that can be used to evaluate ethical perceptions. The tests are those suggested by psychometricians to validate psychological construct. Each test produced results that support the overall validity of the multidimensional measure. All the test criteria seem to support the proposition that the three dimensions—moral equity, relativism, and contractualism—capture a substantial amount of the decision dynamics used by the respondents to make ethical judgments. These dimensions also capture a substantial amount of the variance in the construct "behavorial intention."

The multidimensional measure is a potentially useful tool for directing research on why accountants make certain ethical judgments. Application of the multidimensional measure to a wide variety of situations is likely to produce researchable hypotheses about reasons behind relative judgments. Knowing the grounds on which accountants tend to make ethical judgments can be useful in practical ways; the information can be used to establish ethical norms within an organization and to identify individuals whose behavior may present potential problems.

The ethical perception of a situation is a complex process, involving at least the three dimensions investigated in this study. Different situations may elicit different, and perhaps undiscovered, dimensions, and future research should include attempts to uncover additional dimensions if they exist. Nonetheless, use of the multidimensional measure as a place to begin research may be this study's most beneficial outcome. Continued application of the scale to different business situations with different groups, both accountants and users of accounting services, should improve our understanding of ethical judgments in the practice of accounting. Such an understanding is essential before ethical behavior can be moderated or ethical conduct prescribed.

Appendix

The following four scenarios were used in this study. Each scenario appeared on a separate page followed by brief instructions, a randomized presentation of the scales in figure 1, the univariate ethics measure, and the behavioral intention measure (see fn. b in table 4).

Scenario A

Tom Waterman is a young management accountant at a large, diversified company. After some experience in accounting at headquarters, he has been transferred to one of the company's recently acquired divisions, run by its previous owner and president, Howard Heller. Howard has been retained as vice president of this new division, and Tom is his accountant. With a marketing background and a practice of calling his own shots, Howard seems to play by a different set of rules than those to which Tom is accustomed. So far it is working, as earnings are up and sales projections are high.

The main area of concern to Tom is Howard's expense reports. Howard's boss, the division president, approves the expense reports without review, and expects Tom to check the details and work out any discrepancies with Howard. After a series of large and questionable expense reports, Tom challenges Howard directly about charges to the company for typing that Howard's wife did at home. Although company policy prohibits such charges, Howard's boss again signed off on the expense. Tom feels uncomfortable with this and tells Howard that he is considering taking the matter to the Board Audit Committee for review. Howard reacts sharply, reminding Tom that "the Board will back me anyway" and that Tom's position in the company would be in jeopardy.

ACTION: Tom decides not to report the expense charge to the Audit Committee. Please evaluate this action of Tom Waterman.

Scenario B

Anne Devereaux, company controller, is told by the chief financial officer that in an executive committee meeting the CEO told them that the company "has to meet its earnings forecast, is in need of working capital, and that's final." Unfortunately, Anne does not see how additional working capital can be raised, even through increased borrowing, since income is well below the forecast sent to the bank. Seth suggests that Anne review bad debt expense for possible reduction and holding sales open longer at the end of the month. He also brushes off the management letter request from the outside auditors to write down the spare parts inventory to reflect its "true value."

At home on the weekend, Anne discusses the situation with her husband, Larry, a senior manager of another company in town. "They're asking me to manipulate the books," she says. "On the one hand," she complains, "I'm supposed to be the conscience of the company and on the other, I'm supposed to be absolutely loyal." Larry tells her that companies do this all the time, and when business picks up again she'll be covered. He reminds her how important her salary is to help maintain their comfortable lifestyle, and that she shouldn't do anything drastic that might cause her to lose her job.

ACTION: Anne decides to go along with the suggestions proposed by her boss. Please evaluate this action of Anne Devereaux.

Scenario C

Drew Isler, the plant's chief accountant, is having a friendly conversation with Leo Sullivan, operations manager and old college buddy, and Fred LaPlante, the sales manager. Leo tells Drew that the plant needs a new computer system to increase operating efficiency. Fred interjects that with the increased efficiency and decreased late deliveries their plant will be the top plant next year.

However, Leo wants to bypass the company policy which requires that items greater than $5,000 receive prior Board approval and be capitalized. Leo would prefer to generate purchase orders for each component part of the system, each being under the $5,000 limit, and thereby avoid the approval "hassle." Drew knows this is clearly wrong from a company and an accounting standpoint, and he says so. Nevertheless, he eventually says that he will go along.

Six months later, the new computer system has not lived up to its expectations. Drew indicates to Fred that he is really worried about the problems with the computer, and the auditors will disclose how the purchase was handled in the upcoming visit. Fred acknowledges the situation by saying that production and sales are down and his sales representatives are also upset. Leo wants to correct the problems by upgrading the system (and increasing the expenses), and urges Drew to "hang in there."

ACTION: Feeling certain that the system will fail without the upgrade, Drew agrees to approve the additional expense. Please evaluate this action of Drew Isler.

Scenario D

Paul Tate is the assistant controller at Stern Electronics, a medium-sized manufacturer of electrical equipment. Paul is in his late fifties and plans to retire soon. His daughter has been accepted into medical school, and financial concerns are weighing heavily on his mind. Paul's boss is out of the office recuperating from health problems, and in his absence Paul is making all decisions for the department.

Paul receives a phone call from an old friend requesting a sizable amount of equipment on credit for his new business. Paul is sympathetic but cognizant of the risk of extending credit to a new company, especially under Stern's strict credit policy for such transactions. When Paul mentions this conversation to Warren, the general manager, he is immediately interested. Warren notes that the company needs an additional $250,000 in sales to meet the quarterly budget and, thus, ensure bonuses for management, including Paul.

ACTION: Paul decides to make the sale to his friend's new business. Please evaluate this action of Paul Tate.

References

American Accounting Association Committee on the Future Structure, Content and Scope of Accounting Education. 1986. Future accounting education: Preparing for the expanding profession. *Issues in Accounting Education* 1 (Spring): 168–95.

American Institute of Certified Public Accountants. 1986. *Restructuring Professional Standards to Achieve Professional Excellence in a Changing Environment*. New York: AICPA.

Benston, G. 1982. An analysis of the role of accounting standards for enhancing corporate governance and social responsibility. *Journal of Accounting and Public Policy* 1 (1): 5–17.

Briloff, A. 1972. *Unaccountable Accounting*. New York: Harper & Row.

———. 1976. *More Debits Than Credits*. New York: Harper & Row.

———. 1981. *The Truth About Corporate Accounting*. New York: Harper & Row.

———. 1986a. Accountancy and the public interest. *Advances in Public Interest Accounting* 1: 1–14.

———. 1986b. Standards without standards/principles without principles/fairness without fairness. *Advances in Accounting* 3: 25–50.

Brown, M., K. Feldman, S. Schwartz, and A. Heingartner. 1969. Some personality correlates of conduct in two situations of moral conflict. *Journal of Personality* 37 (March): 41–57.

Burchell, S., C. Clubb, A. Hopwood, J. Hughes, and J. Nahapiet. 1980. The roles of accounting in organizations and society. *Accounting, Organizations and Society* 5 (1): 5–27.

Burton, R. 1963. Generality of honesty reconsidered. *Psychological Review* 70 (6): 481–99.

Carey, J. 1946. *Professional Ethics of Public Accounting*. New York: AICPA.

———. 1947. The realities of professional ethics. *The Accounting Review* 22 (April): 119–23.

Colignon, R., and M. Covaleski. 1991. A Weberian framework in the study of accounting. *Accounting, Organizations and Society* 16 (2): 141–57.

Davis, R. 1984. Ethical behavior reexamined. *The CPA Journal* 54 (December): 32–36.

Ferrell, O., and L. Gresham. 1985. A contingency framework for understanding ethical decision-making in marketing. *Journal of Marketing* 49 (Summer): 87–96.

Finn, D., L. Chonko, and S. Hunt. 1988. Ethical problems in public accounting: The view from the top. *Journal of Business Ethics* 7 (August): 605–15.

Fishbein, M., and I. Ajzen. 1975. *Belief, Attitude, Intention and Behavior: An Introduction to Theory and Research*. Reading, MA: Addison-Wesley.

Fletcher, J. 1966. *Situation Ethics.* London: SCM Press.

Gaa, J. 1986. User primacy in corporate financial reporting: A social contract approach. *The Accounting Review* 61 (July): 435–54.

———. 1990. A game-theoretic analysis of professional rights and responsibilities. *Journal of Business Ethics* 9 (March): 159–69.

Graber, D. 1979. Ethics enforcement—how effective? *The CPA Journal* 49 (September): 11–17.

Gray, S. 1988. Towards a theory of cultural influence on the development of accounting systems internationally. *Abacus* 24 (1): 1–15.

Hartshorne, H., and M. May. 1928. Studies in the nature of character. *Studies in Deceit*, vol. 1. New York: Macmillan.

Hoffman, M. 1977. Moral internalization: Current theory and research. In *Advances in Experimental Social Psychology*, vol. 10, edited by L. Berkowitz. New York: Academic Press.

Hunt, S., and S. Vitell. 1986. A general theory of marketing ethics. *Journal of Macromarketing* 6 (Spring): 5–16.

Lehman, C. 1988. Accounting ethics: Surviving survival of the fittest. *Advances in Public Interest Accounting* 2: 71–82.

———, and T. Tinker. 1987. The "real" cultural significance of accounts. *Accounting, Organizations and Society* 12 (5): 503–22.

Lev, B. 1988. Toward a theory of equitable and efficient accounting policy. *The Accounting Review* 63 (January): 1–22.

Loeb, S. 1971. A survey of ethical behavior in the accounting profession. *Journal of Accounting Research* 9 (Autumn): 287–306.

———. 1972. Enforcement of the code of ethics: A survey. *The Accounting Review* 47 (January): 1–10.

———. 1988. Teaching students accounting ethics: Some crucial issues. *Issues in Accounting Education* 3 (Fall): 316–29.

Lowe, H. 1987. Ethics in our 100-year history. *Journal of Accountancy* 163 (May): 78–87.

Magill, H., and G. Previts. 1991. *CPA Professional Responsibilities: An Introduction.* Cincinnati: South-Western Publishing Co.

Mautz, R. 1984. Self-regulation: Criticisms and a response. *Journal of Accountancy* 157 (April): 56–66.

———. 1988. Public accounting: Which kind of professionalism? *Accounting Horizons* 2 (September): 121–25.

Merino, B., and M. Neimark. 1982. Disclosure regulation and public policy: A sociohistorical reappraisal. *Journal of Accounting and Public Policy* 1 (1): 33–57.

National Commission on Fraudulent Financial Reporting. 1987. *Report of the National Commission on Fraudulent Financial Reporting.* Washington, D.C.: United States Government Printing Office.

Nelson, E., R. Grinder, and M. Mutterer. 1969. Sources of variance in behavioral measures of honesty in temptation situations: Methodological analyses. *Developmental Psychology* 1 (3): 265–79.

Noreen, E. 1988. The economics of ethics: A new perspective on agency theory. *Accounting, Organizations and Society* 13 (4): 359–69.

Nunnally, J. 1967. *Psychometric Methods.* New York: McGraw-Hill.

———. 1978. *Psychometric Methods.* 2d ed. New York: McGraw-Hill.

Parker, L. 1987. An historical analysis of ethical pronouncements and debate in the Australian accounting profession. *Abacus* 23 (2): 122–40.

Parker, R. 1986. The development of the accountancy profession in Britain to the early twentieth century. Monograph 5. The Academy of Accounting Historians: School of Accountancy, James Madison University, Harrisonburg, Virginia.

Reidenbach, R., and D. Robin. 1990. Toward the development of a multidimensional scale for improving evaluations of business ethics. *Journal of Business Ethics* 9 (August): 639–53.

Richardson, A. 1931. *The Ethics of a Profession.* New York: The Century Co.

Shaub, M. 1988. Restructuring the code of professional ethics: A review of the Anderson committee report and its implications. *Accounting Horizons* 2 (December): 89–97.

Solomons, D. 1991. Accounting and social change: A neutralist view. *Accounting, Organizations and Society* 16 (3): 287–95.

Tinker, T. 1985. *Paper Prophets: A Social Critique of Accounting.* New York: Praeger.

——. 1988. Panglossian accounting theories: The science of apologizing in style. *Accounting, Organizations and Society* 13 (2): 165–89.

——, B. Merino, and M. Neimark. 1982. The normative origins of positive theories: Ideology and accounting thought. *Accounting, Organizations and Society* 7 (2): 167–200.

Touche Ross. 1988. Ethics in business: An opinion survey of key business leaders on ethical standards and behavior. New York: Touche Ross.

Watts, R., and J. Zimmerman. 1978. Towards a positive theory of the determination of accounting standards. *The Accounting Review* 53 (January): 112–34.

——. 1986. *Positive Accounting Theory.* New York: Prentice-Hall.

Williams, P. 1987. The legitimate concern with fairness. *Accounting, Organizations and Society* 12 (2): 169–89.

Woelfel, C. 1986. Standards of ethical conduct for management accountants. *Journal of Business Ethics* 5 (May): 365–71.

THE ACCOUNTING REVIEW
Vol. 67, No. 2
April 1992
pp. 303–318

Risk Preferences in Participative Budgeting

D. C. Kim
Donge-Guk University (Korea)
and Louisiana State University

SYNOPSIS AND INTRODUCTION: This study examines participative budgeting[1] in the context of the psychology of risk. As Young (1985) and Waller (1988) report, there is some preliminary evidence that risk-averse workers create more budgetary slack than risk-neutral ones. They show that "truth inducing incentive schemes" (e.g., Soviet incentive schemes; see Weitzman 1976) reduce budgetary slack for risk-neutral subjects but not for risk-averse subjects. If this is true, it means that resource allocations within organizations are mediated by perceptions of risk.

Young (1985) and Waller (1988) define risk preference as a dispositional variable,[2] which presumes that it is a stable personal trait, a latent variable, traditionally inferred from observed behavior of risk propensity in such settings like lotteries. This study tests whether risk preferences are domain-specific; that is, latent risk preferences translate into differing manifest risk preferences according to the context. Domain-specific risk preference can be understood as a manifest psychological variable that may well be the result of the combination of latent risk propensity and the situation. Kahneman and Tversky's prospect theory (1979) suggests that manifest risk preferences depend upon whether the subject frames his or

---

[1] Participative budgeting refers to a specific managerial style in which subordinates are permitted to participate with their superiors in setting their performance budget (e.g., next period's budgeted production or budgeted sales; see Locke and Schweiger 1979). Participation is defined here as a budgeting regime where subordinates select performance goals jointly with their superiors (Young 1985).

[2] Young defines a risk-averse subordinate as one whose expected utility function for money is concave (Young 1985, 830) and measures the disposition with a conventional two-outcome lottery in a context-free scenario. Waller employs the induced utility function procedure of Berg et al. (1986). Thus, both studies are within the framework of the traditional expected utility paradigm which does not consider situation-dependent risk preferences.

This paper is based on my dissertation at the University of Iowa. I thank the members of my dissertation committee: Edward Arrington, Irwin Levin, Jere Francis, and Rick Tubbs. Special acknowledgement is due to my dissertation chairman, Albert Schepanski. I would also like to acknowledge the helpful comments on earlier versions of this paper by accounting workshop participants at the University of Iowa and Louisiana State University.

Submitted August 1989.
Accepted October 1991.

her task in the context of gain or loss prospects, where gains and losses are defined in relation to a neutral reference point. If risk preferences are domain-specific, then past studies' suggestion that incentive schemes should be designed in consideration of *dispositional*, or latent, risk preferences needs to be reexamined (Waller 1988; Kaplan 1982).

The question before this study is: If subordinates are influenced by prior period performance in setting current period budgets for themselves, will that influence take the form predicted by prospect theory and thus lead to riskier preferences (tight budgets) when the subordinates perceive themselves in a losing situation? This is an important question considering Young's (1985) alluding to the possibility that inducing subordinates to less risk-averse behavior may be a way to favor tight budgetary standards and reduce slack. This prediction is consistent with prospect theory's implication that losers who are slow to adjust their reference point act in a more risk-seeking manner.[3] Thus, the induction of losing prospects might be a way to minimize budgetary slack.

An additional concern in this study is to jointly test domain-specific risk preferences and dispositions toward risk as influences over budgetary decisions. Whereas prospect theory explains risk preferences as domain-specific contingencies, other theories construe risk preferences as dispositional. It is likely that both domain-specific and dispositional factors influence budgetary decisions. This study deploys a conventional lottery procedure to elicit and test dispositions toward risk.

An experiment simulating the public accountants' budgeting of billable hours was designed to test the hypothesis that subject preference for tight or safe budget behavior depends on the performance of coworkers and domain-specific risk preferences. The hypotheses were tested in an experiment employing 81 students. The results generally support the view that subordinates' risk preferences are influenced by a situation-dependent variable. The reversal of risk preferences around a neutral reference point is statistically significant for both dispositionally risk-averse and dispositionally risk-seeking subjects. The dispositional variable also contributes to the explanation of variations in subjects' manifest risk preferences. Thus the propensity to induce budgetary slack seems to be a joint function of situations and dispositions.

Key Words: *Participative budgeting, Risk preference, Prospect theory, Budgeting.*

Data Availability: *The experimental data in this study are available from the author upon request.*

---

[3] Tversky and Kahneman (1981) proposed that people generally evaluate alternatives in terms of either a minimal account or a more inclusive mental account. The minimal account includes only the direct consequences of a decision at hand. The inclusive mental account refers to situations in which the outcomes of an action affect the balance in a mental account that was previously set up by related actions, as in the case of a bettor at the horse track who views the last race in the context of earlier races in the same day. According to prospect theory, if one is in a losing condition, the more inclusive account should lead to more risk-seeking behavior than the minimal account.

THE first section below presents the theoretical background of this study with a brief review of relevant literature. This is followed by the research hypotheses in the second section. The third section outlines the experimental task and procedures used to test the hypotheses. The results of the experiment are presented in the fourth section. The final section discusses the implications of the major findings of this study and makes suggestions for future research.

I. Background

Domain-specific Risk Preferences in Prospect Theory

Tversky and Kahneman (1981) proposed prospect theory as an alternative theory of decision-making behavior under risk. They present empirical findings that choice behavior consistently violates certain axioms of expected utility theory. After reviewing recent evidence concerning the expected utility model, Schoemaker (1982) concluded that at the individual level most of the empirical evidence is difficult to reconcile with the principle of expected utility maximization. As such, the descriptive validity of that principle at the individual level remains in question.

Prospect theory is one way to enhance this line of research inasmuch as it contains some identifiable differences from expected utility theory, differences that are empirically testable. The major differences between prospect theory and expected utility theory concern prospect theory's value function and its decision-weighting function. The value function, which is the Kahneman-Tversky version of a utility function, is assumed to be concave for gains, convex for losses. The decision-weighting function reflects the impact of events on the overall attractiveness of gambles. From these functions, Kahneman and Tversky point out that a positive or negative translation of regular prospects may raise or lower the reference point. Likewise, a change in the reference point may change the preference ordering among a set of choice problems.

Prospect theory has emphasized the importance of a reference point in determining preferences and choices among prospects. Reference points represent variations on the concept of an aspiration level. Prospect theory suggests that the location of the reference point and the way in which decision problems are coded and edited are critical factors in the analysis of decision-making behavior.

The theory posits a reflection effect that refers to the reversal of preferences around a neutral reference point (no change in wealth). The reflection effect is implied by the value function that is concave on the positive domain and convex on the negative domain.

Prospect theory proposes that people evaluate outcomes as deviations from a reference point. Given the concavity in the gain domain and convexity in the loss domain, the value function in prospect theory implies risk-aversion if outcomes are evaluated as gains and risk-seeking if outcomes are evaluated as losses in relation to an individual's reference point. The reflection effect and framing effects result from this evaluation process. Since expected utility theory defines individuals' utility over final wealth states, the framing of outcomes as gains or losses would be presumed irrelevant as long as the final wealth states associated with alternatives are identical. Therein lies the major difference between the two theories.

Broadly speaking, the theoretical perspective of Kahneman and Tversky and related work suggests the following descriptive set of assumptions about risk attitudes:

(1) the utility (or value) function is defined in terms of levels and changes in current, rather than terminal, wealth; (2) changes in wealth are evaluated relative to a reference point; and (3) risk preference is conditioned by a reference point (for a detailed discussion, see Kahneman and Tversky 1979; Libby and Fishburn 1977; Payne et al. 1980).

Empirical Evidence of Domain-specific Risk Preferences

The reflection effect has often been tested using choices of gambles. Payne et al. (1980, 1981) presented to subjects pairs of gambles with varying reference points manipulated by adding or subtracting a constant amount from all outcomes. Their results generally supported the reflection effect.[4] However, other studies report evidence that questions the generality of the reflection effect (Hershey and Schoemaker 1980; Schneider and Lopes 1986; Cohen et al. 1987).

Leopard (1978) found that most subjects took more risk when they were "behind" than when they were "ahead" in repeated gamble experiments. The subjects apparently risked more in an effort to catch up. It appears that subjects were unable to adjust their reference points to their current wealth position. In such instances, mental accruals of past outcomes seem to be more salient than the independence of plays.

Evidence consistent with the reflection effect has emerged in both laboratory experiments and field settings. Metzger (1985) found that bettors at the horse track are more speculative in later races than in early races. This could be accounted for by the notion that the reference point at the beginning of a racing day (e.g., break-even) is not adjusted to subsequent losses. Similar results are reported in capital markets. Lakonishok and Smidt (1986) and Shefrin and Statman (1985) provide evidence that investors have a tendency to sell winning stocks too early (concave value function in the gain domain) and to hold losing stocks too long (convex value function in the loss domain).

Several accounting studies investigated framing effects in a tax context. Such effects are closely related to the reflection effect in that the framing effect depends upon prospect theory's reflection effect.[5] The evidence is generally mixed. While Chang et al. (1987) and Schepanski and Kelsey (1990) found results consistent with prospect theory's implications, Schadewald (1989) observed no significant framing effects in tax reporting decisions.

Recent research effort refines both the reflection effect and decision-weighting functions in prospect theory (Hogarth and Einhorn 1990). Schoemaker (1990) also provides some support to the notion of domain-specific risk preferences in prospect theory. Several studies extend the implications of prospect theory into a multi-period context (Loewenstein 1988; Thaler and Johnson 1990; Loewenstein and Prelec 1991). Rachlin (1990) also provides an interesting extension of prospect theory into multi-period situations involving a string of losses.

In summary, prospect theory's assumption of domain-specific effects on risk preferences is an important concept in current research on decision-making under uncertainty, although not all of the empirical evidence is consistent with it. Given the initial evidence that dispositional risk attitude is a determinant of budgetary slack (Young 1985; Waller 1988), this study's major contribution is to introduce prospect theory's domain-specific risk preference in a budgetary setting.

---

[4] Petty and Scott (1981) conducted a survey of capital budgeting practices of *Fortune 500* firms and found that the idea of the target return is widely used as a reference point in business decision making.

[5] The framing effect posits that objectively identical problems (in terms of final wealth states) are perceived differently depending on how those problems are described. Different framings of a problem often induce different behaviors. A framing effect would be expected if the reflection effect holds. If the reflection effect does not hold, however, there would be no reason to expect choices to differ between different framing conditions.

II. Research Hypotheses

The most relevant feature of prospect theory for this study is domain-specific risk preferences: risk-averse behavior above a reference point and risk-seeking behavior below. The location of a reference point is critical in testing domain-specific risk preferences and often difficult to operationalize. To overcome that difficulty, Fishburn's (1977) suggestion that people perceive the average performance of others in the same task as their reference point is adopted in this study (see also Payne et al. 1980, 1981).

Relying on such work, this study assumes that a subordinate perceives the average performance of fellow workers in a comparable position as the reference point. Competition in a business environment and an individual's self-esteem support the validity of this assumption inasmuch as both competition and self-esteem can be derived through comparison of the self with others. For example, an entrepreneur who is weathering a slump with greater success than his competitors may interpret a small loss as a gain, relative to the larger loss he had reason to expect (Kahneman and Tversky 1979, 286).

Using Kahneman and Tversky's notation of choice problems, let $(x, p; y, 1-p)$ represent a prospect which yields x with probability of p and y with probability of $1-p$. The typical participative budgeting problem can then be represented as follows:

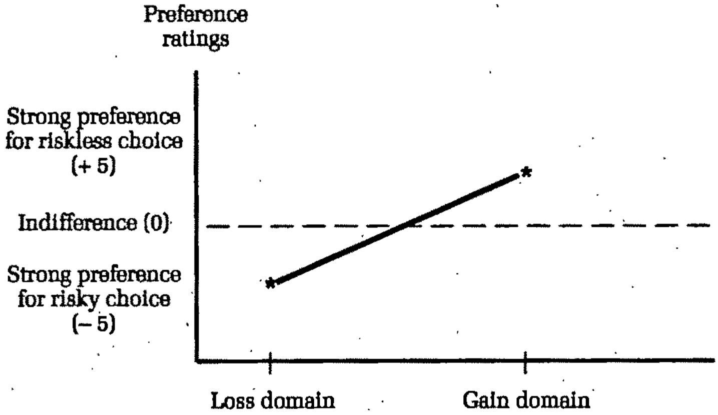
$$\begin{aligned} &\text{A. } (M1, p; M4, 1-p) \\ &\text{B. } (M2, 1; M3, 0), \end{aligned} \tag{1}$$

where $M1$ is the outcome if a tight budget is chosen and achieved; $M2$ is the outcome if a safe budget is chosen and achieved; M3 is the outcome if a safe budget is chosen but not met; $M4$ is the outcome if a tight budget is chosen but not met; and, p is the probability that a subordinate will achieve a tight budget. Thus budgeted performance in alternative A represents a risky choice, whereas budgeted performance in alternative B represents a riskless choice with a certain outcome. The outcomes of a tight budget ($M1$ and $M4$) correspond to gain and loss domains, respectively, and the outcome of a safe budget ($M2$) is the certainty equivalence of the outcomes of a tight budget (M1 and M4). A tight budget (A) has a greater outcome variance than a safe budget (B). Achieving a tight budget (M1) rewards more than achieving a safe budget (M2), whereas failing a tight budget (M4) is more detrimental to performance than achieving a safe budget (M2). That is, $M1$ is greater than M2 and M4 is smaller than M2 ($M1>M2>M4$). The relationships among M1, $M2$, and M4 are critical, whereas the outcome of failing a safe budget is not important in that the safe budget is a budget with excessive budgetary slack, a budget which is certain to be met.

Hypothesis H1

If the subordinate views the average performance of his or her fellow workers as a reference point, then the subordinate's risk preferences would depend, in part, on his or her current status relative to others; that is, whether performance is below or above

Figure 1
Predictions of Hypothesis H1



the average performance, with average performance operationalized as a zero budget variance. Domain-specific risk preferences suggested by prospect theory predict risk-seeking in a loss domain and risk-aversion in a gain domain. If the subordinate is below the average performance of fellow workers, he or she would choose a tight budget if the reward associated with meeting a tight budget is sufficiently large to make up for the underachievement in preceding periods. In other words, the subordinate would take a risky action if the difference between $M1$ and $M2$ is equal to or greater than the magnitude of unfavorable discrepancy between his or her performance and the average performance in a prior period (eq. [1]). If the subordinate has a favorable variance (a gain position), he or she would choose a safe budget since the concave value function implies that that the subjective value associated with a safe budget is greater than the subjective value of a tight budget. Hypothesis H1 then follows from the above predictions (fig. 1). Empirical results will support the first hypothesis if the subjects' mean response in the loss condition is significantly skewed toward the tight budget while the mean response in the gain condition is significantly skewed toward the safe budget.

H1: The subordinate is more likely to prefer a tight budget when below, and a safe budget when above the average performance of fellow workers.

Hypothesis H2

Past studies in participative budgeting (Young 1985; Waller 1988) treated risk preference as a dispositional characteristic of the individual; that is, as a latent and invariant attribute of personality. With such a view, dispositionally risk-averse subjects would always prefer safe budgets to tight budgets, whereas dispositionally risk-seeking subjects would always prefer tight budgets to safe budgets regardless of their performance relative to the average performance of other workers.

This study tests prospect theory's proposition that the antecedent of a subordinate's preference for a tight or a safe budget is a domain-specific variable. Given the finding by Young (1985) and Waller (1988) that dispositional risk attitude is a determinant of budgetary slack, a question arises: Do both groups of subjects whose attitudes toward risk are dispositionally different (risk-averse and risk-seeking) exhibit different preferences for tight as opposed to safe budgets depending on the average performance of fellow workers?

Prospect theory proposes that a domain defined by a reference point is the *only* relevant variable in explaining risk preferences. That is, the shape of the value function is presumed determined by a situation relative to a reference point but *not* by disposition. This is a very strong claim inasmuch as it is insensitive to the tenets of expected utility theory. Apart from Lopes (1984) and Schneider and Lopes (1986), extant research does not attempt to jointly test both situational effects (as posited by prospect theory) and dispositional ones (as posited by expected utility theory). Lopes (1984) suggests that dispositionally risk-averse and risk-seeking subjects differ in judgments of riskiness—the two groups of subjects may differ in whether they pay most attention to the worst or the best outcomes in risky distributions. Schneider and Lopes (1986) find that prospect theory's value function describes the preferences of dispositionally risk-averse subjects for gains and dispositionally risk-seeking subjects for losses, although the reflection effect is weak and irregular except for lotteries with riskless components. These studies provide limited implications for the situational effects in that situational variables are manipulated with context-free lotteries involving positive or negative outcomes.

This study makes some effort to jointly test the two theories in an accounting setting, especially since previous studies show that dispositions toward risk influence budgetary decisions. Specifically, the study examines whether *both* dispositionally risk-seeking and dispositionally risk-averse subjects will prefer safe budgets if they are above the average performance of fellow workers and tight budgets if they are below the average performance of fellow workers. The second hypothesis will be tested by analyzing mean responses of dispositionally risk-seeking and dispositionally risk-averse groups separately.

H2: Both dispositionally risk-seeking and dispositionally risk-averse groups exhibit domain-specific risk preferences.

III. Method

Design

The experimental design is a 2×2 factorial. The factors are the subject's status relative to average performance (below average vs. above average) and the subject's dispositional risk attitude (risk-averse vs. risk-seeking). The situational factor was manipulated as a between-subjects factor, while the disposition toward risk variable was measured using a conventional lottery. The subjects' risk attitudes (dispositions) were measured first. Subjects were then randomly assigned to a gain condition (above average performance condition) or a loss condition (below average performance condition). Table 1 displays the experimental design and predicted directions of mean responses in each condition.

Subjects

Subjects were 81 undergraduate seniors majoring in accounting at a Big Ten university. The experiment was introduced as a study of public accountants' decision

Table 1
Experimental Design

	Situation	
Disposition	*Loss Domain*	*Gain Domain*
Risk-seeking	Cell A*	B
Risk-averse	C	D

Comparison of Mean Preference Ratings

Hypothesis H1	$(A+C)/2<0$ and $(B+D)/2>0$
Hypothesis H2	$A<0$ & $B>0$ and $C<0$ & $B>0$

* Cell entries are mean preference ratings.

behavior in the context of budgeted billable hours. Budgeted billable hours was chosen for the experimental task since subjects were expected to relate to it because most of them will soon begin careers as accountants. Furthermore, an experimental task that is salient and personally relevant is more likely to lead subjects to devote more attention to the task (Deci 1975; Fischhoff 1982). Personal relevance and salience are particularly important in studies of judgment rather than performance since it is difficult to utilize monetary payments as incentives in studies of judgment. Accounting seniors are generally concerned about their future careers in public accounting, and the hypothetical cases of budgeting and performance evaluation in a public accounting firm would be expected to be salient experimental stimuli to them. The questionnaires were reviewed for contextual considerations by a former partner in a large public accounting firm. The experiment took approximately 20 minutes to complete.

Procedure

Dispositional Risk Attitude Measurement. The subjects received risk attitude measurement questionnaires first. The subjects' risk attitudes were measured in a conventional way. Each subject indicated his or her willingness to play a gamble similar to gambles used by Kahneman and Tversky (1979).[6] The gamble was a mixed gamble involving a choice between a 0.5 chance of winning $10 and a 0.5 chance of losing $10 or, alternatively, a choice of no play. The subjects' response scale ranged from 1 (extreme willingness) to 9 (extreme unwillingness). The mean of the subjects' dispositional risk attitudes was 5.3, which suggests, as one would perhaps expect, an overall indifference for playing the gamble.[7]

[6] The hypothetical gamble was chosen rather than a real gamble since the subjects playing a lottery with an initial endowment of $10 may perceive the gamble as a choice involving gains for which prospect theory predicts risk aversion. If so, the lottery becomes a gamble of (20, 0.5; 0, 0.5), as opposed to a gamble of (10, 0.5; −10, 0.5), in Kahneman and Tversky's notation (1979). Given the short duration of the experiments, the subject's reference point in the lottery is likely to be his or her net wealth before the endowment.

[7] To control any possible confounding effect due to the response scale, one half of the subjects received the opposite order of willingness scale which ranges from 1 (extreme unwillingness) to 9 (extreme willingness). The different orderings of risk attitude measurement scales did not produce any difference in responses at conventional probability levels.

Based on the elicitation, subjects were classified as risk-seeking or risk-averse according to their dispositions toward lotteries. A subject was classified as risk-averse if his or her response was above 5 (indifference) and risk-seeking if equal to or below 5. This procedure resulted in an approximately equal number of subjects in each dispositional risk attitude group; the risk-seeking group included 40 subjects and the risk-averse group included 41 subjects.

Main Experimental Sessions. Subjects then read a case about a hypothetical audit engagement. Each subject assumed the role of an entry level accountant in a public accounting firm with no work history. This was done to control any confounding effect of work history on the budgetary choice, though there is little reason to expect such an effect given that subjects had very little work history. The budgeted billable hours and the variance between the budgeted and the actual billable hours were described briefly. The accounting firm's performance evaluation scheme was described as involving the use of the time budget variance as a measure of the accountant's performance. The subjects were instructed that a zero variance was an average performance of all fellow workers in the firm.

The subjects' situational variable was a loss or a gain dimension, depending upon the experimental condition. The loss condition was described as a tentative unfavorable variance in billable hours, and the gain condition was described as a tentative favorable variance in billable hours (refer to appendix A). The subjects were told that they were in the final stage of the audit and that their budgetary choices in this stage and the resulting outcomes would determine the total variances of the audit and therefore their overall performance.

Subjects were then asked to choose between a tight budget and a safe budget and to indicate the strength of their preference. The subjects' responses in this stage are the primary focus of this study. In the loss condition, the tight budget has two possible outcomes—a 0.5 chance to incur a slightly favorable variance for the entire audit engagement, and a 0.5 chance to incur another unfavorable variance in the coming month, an outcome that would result in a large unfavorable variance for the entire audit engagement. The safe budget has one certain outcome in which no variance in the current period is incurred. However, choosing the safe budget would make it impossible to overcome the unfavorable variance that was accrued in the prior month. As a result, the subject who chooses a safe budget would guarantee the overall variance for the entire audit engagement to be unfavorable. The gain condition is similar to the loss condition except that the subject has a favorable variance prior to the completion of the audit engagement. In this gain setting, choosing the safe budget would result in a slightly favorable variance for the entire audit, whereas choosing the tight budget would yield a 0.5 chance of a large favorable variance and a 0.5 chance of a slightly unfavorable variance for the entire audit. In both conditions, the expected values of tight and safe budgets are the same.

Subjects' responses are recorded by combining the response to a forced choice with a preference rating into a single measure. The choice of a tight (risky) budget is recorded as "$-$" and the choice of a safe (riskless) budget is recorded as "$+$". Thus, combined responses vary from -5 (strong preference for a risky choice) to a $+5$ (strong preference for a riskless choice). If combined responses are close to zero (either -1 or $+1$), for example, they are interpreted as slight preferences. The analysis uses the combined measure as the dependent variable. The forced choice represents the subject's binary choice of either type of budget. The subject's preference rating of a selected bud-

Table 2

Profile of Subjects

Demographic Variables	
Mean age	22.3
Sex:	
Female	45 subjects
Male	36 subjects
Class standing:	
Accounting Senior	79 subjects
Accounting graduated	2 subjects
Mean semester hours in cost accounting	5.7 semester hours
Mean semster hours in auditing	3.0 semester hours
Career objective:	
Public accounting	54 subjects (66.7%)
Accounting-related jobs	19 subjects (23.5%)
Nonaccounting job	8 subjects (9.9%)
Job experience:	
No experience	50 subjects
Part-time experience	28 subjects
Full-time experience	3 subjects

get represents the strength of his or her preference. Thus the combined measure of choice and preference rating represents a finer measure of the subject's preference. The major conclusions apply to both the forced choice responses and to the combined measures (Levin et al. 1987).

Manipulation Check and Profile of Subjects. After the main experimental trials were completed, the subjects responded to a manipulation check question. That question elicited a self-report on their performance prior to the final stage of the audit relative to the average performance of other workers. The subjects' responses could range from 1 (below average), to 2 (average), and to 3 (above average). The mean response to the question was 1.37 in the loss condition and 2.65 in the gain condition. These were significantly different ($t=11.9$, $p<0.001$). That is, the status variable was manipulated as intended.

Table 2 shows a profile of subjects. Two thirds of the subjects indicated the practice of public accounting as their immediate career objective. Only eight subjects (10 percent) indicated that their career objective was other than accounting-related jobs. Thus, as expected, the subjects were a relatively homogeneous group with respect to future career objectives. Most subjects had six semester credits of cost accounting and three semester credits of auditing. Thirty-eight percent of the subjects had work experience, most of which involved summer internships in accounting firms, part-time cashiers, and other limited experience.[8]

---

[8] The demographic variables in table 2 were tested for any possible significance in mean preference ratings. None of them were significant at conventional probability levels.

Table 3
ANOVA Results*

	df	*Sum of Square*	*Mean S.S.*	F	*Probability*
Situational variable (*A*)	1	243.2	243.2	28.09	0.0001
Dispositional variable (*B*)	1	57.8	57.8	6.64	0.0119
Interaction (*A***B*)	1	4.5	4.5	0.52	0.4722
Error	77	670.8	8.4		
Total	80	976.3			

* $R^2 = 0.31$.

IV. Results

Overview

Overall, the results of the experiment supported the predictions of this study derived from prospect theory *as well as* the conclusions that dispositions toward risk do condition budgetary choice. The overall ANOVA results showed that both the situational variable and the dispositional variable were statistically significant ($F = 28.09$, $p < 0.001$, and $F = 6.64$, $p < .01$, respectively). The interaction between the two independent variables was not statistically significant ($F = 0.52$, $p < 0.47$). The results of the ANOVA test are summarized in table 3. The purpose of the ANOVA test is to check the overall significance of the two independent variables in explaining variations in the subjects' budgetary choices.[9] The ANOVA test was conducted as a mixed model with the situational variable treated as a fixed effect and the dispositional variable treated as a random effect. Specific predictions were further tested by examining the signs and the magnitudes of mean preference ratings as displayed in table 1. Tests of the two hypotheses were one-tailed *t*-tests.

Hypothesis H1

The first hypothesis posits that the subordinate is more likely to prefer a tight budget when below, and a safe budget when above the average performance of other workers. The mean preference ratings of the subjects were -0.36 in the loss condition and 3.10 in the gain condition. This difference is significant at $p < 0.001$ ($t = 5.14$). The signs of the mean preference ratings are in the predicted direction. However, the mean preference rating in the loss condition was not significantly less than zero ($t = 0.68$, $p < 0.26$). That is, the subordinates preferred the safe budget over the tight budget when they were performing above the average performance of fellow workers, but they were apparently indifferent between the tight budget and the safe budget when they were below the average performance of fellow workers. Although the mean preference

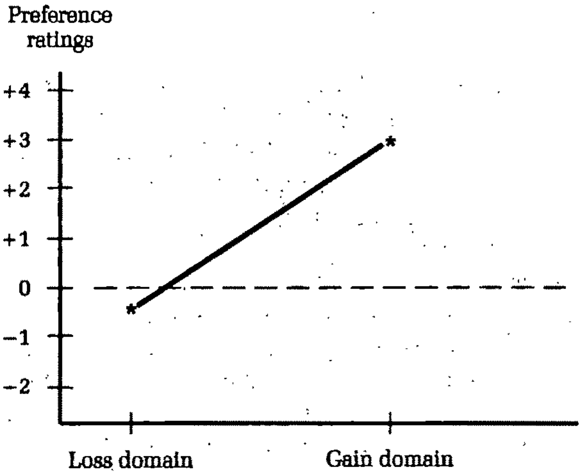
[9] A Chow test was conducted to examine the incremental explanatory power of the situational variable over the dispositional variable. The result was significant ($F = 7.50$, $p < .0001$).

Table 4
Mean Preference Ratings

	Preference Ratings	
	Loss Domain	*Gain Domain*
All Subjects	−.36	3.10*
Risk Attitude		
Risk-seeking subjects	−1.53*	2.52*
Risk-averse subjects	0.64	3.74*

* Significantly different from 0 at $p<0.05$.

Figure 2
Preference Ratings of All Subjects

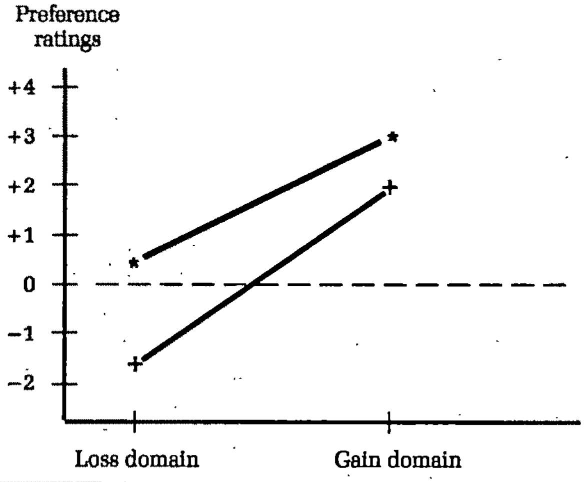


rating in the loss condition is not consistent with risk-seeking behavior for losses, the results provide some support for the reflection effect in that significantly different preferences were observed across experimental conditions. The data are summarized in table 4 and figure 2.

Hypothesis H2

The second hypothesis predicts that the reflection effect holds for both dispositionally risk-seeking and dispositionally risk-averse subjects. Mean preference ratings of dispositionally risk-seeking subjects were −1.53 in the loss condition and 2.52 in the gain condition. The two mean values were significantly different ($t=4.02$, $p<0.001$). The mean rating in the loss domain was significantly less than zero ($p<0.05$; $t=1.94$). Likewise, the mean rating in the gain domain was significantly greater than zero ($p<0.001$; $t=3.95$). (See table 4 and fig. 3.)

Figure 3
Preference Ratings of Risk-averse vs. Risk-seeking Groups



* Mean preference ratings of the dispositionally risk-averse subjects.
+ Mean preference ratings of the dispositionally risk-seeking subjects.

From these results, it appears that the reflection effect was more pronounced among the dispositionally risk-seeking subjects than among the dispositionally risk-averse subjects. Mean preference ratings for the dispositionally risk-averse subjects were 0.64 in the loss condition, and 3.74 in the gain condition (table 4 and fig. 3). The mean rating in the gain condition was significantly greater than zero ($t=8.4$, $p<0.001$), whereas it was not significantly less than zero in the loss condition ($t=0.92$, $p<37$). Since the reflection effect predicts a negative mean preference rating in the loss condition, the observed mean rating in the loss condition does not support the prediction that the dispositionally risk-averse subjects would be more likely to prefer a tight budget when below the average performance of other workers. However, even these results can be interpreted as weak support for the reflection effect in that the dispositionally risk-averse subjects' mean ratings were significantly lower in the loss than in the gain condition ($t=3.65$, $p<0.0005$). This suggests that the dispositionally risk-averse subjects showed a significantly different budgetary choice behavior, depending on their performance relative to the average performance of fellow workers.

Tests of hypothesis H2 provided additional support for hypothesis H1. That is, significantly different risk preferences were observed not only for all the subjects on average but also across the subjects who differed in their dispositional attitude toward risk. This is an important extension of past findings that the dispositional risk attitudes measured by context-free choice problems can have a significant impact on subordinates' self-selection of performance budgets. These results also lend support to the notion that both disposition and situation condition risk preferences and, through those preferences, budgetary choices.

The mean preference rating of dispositionally risk-seeking subjects was significantly lower than that of dispositionally risk-averse subjects (0.6 vs. 2.07) across the two domains ($t = 1.93$, $p < 0.01$). This is consistent with the claims made by Young (1985) and Waller (1988) that different budgetary decisions result from dispositionally different risk attitudes.

V. Discussion and Concluding Remarks

This study addressed two specific issues: does context influence risk preferences; and, if so, does context provide incremental explanatory power over and above dispositions toward risk? The results generally support positive answers to both questions. The most significant finding is that risk preference is a function of the combined effect of context and disposition. The empirical test in this study can be viewed as a joint test of the reference point concept and domain-specific risk preferences. The findings are consistent with the view that subordinates set their reference point (or aspiration level) to the average performance of other workers as well as the view that subordinates are risk-averse for outcomes above the reference point and risk-seeking for outcomes below it.

The evidence presented in this paper raises stimulating questions for future research. However, this study has several limitations. First, the tendency to prefer a tight budget may apply only to a subordinate whose below-average performance is due to exogenous variables; for example, chance-related events or externally imposed constraints. Consistent below-average performance due to inherently low ability may lead to the feeling of "learned helplessness" (Seligman 1975) and thus may not lead such subordinates to lower budgetary slack. Second, this study employed a measure of dispositional risk attitude similar to past studies (Young 1985; Schneider and Lopes 1986). It is possible that this crude measure of risk attitude may have contributed to the limited effect of the dispositional variable on subject responses.

Appendix A

Experimental Instrument for the Loss Condition

You are currently involved in conducting the audit for Tri-State Company, one of the largest clients in your office. The *total* time budget you established to complete the entire audit for this client at the beginning of the audit was 300 hours. The audit is nearing its completion. Unfortunately, you have experienced difficulties in the course of this audit. You budgeted 200 hours for the audit work that you have done so far. You have used 220 hours. As such, you have a tentative *unfavorable variance of 20 hours.*

The work that you still must do to complete this audit was originally budgeted to take 100 hours. You have the task of revising this original budget. This is the final stage of the audit. At the end of this stage, your superior will assess your total budget variance and review your performance.

You must choose Option A or Option B below.

Option A

Set the monthly time budget at 70 hours. If you elect this monthly budget, you must work more quickly than usual in completing the remaining audit work. One of the following two outcomes will then occur.

Outcome 1. You will be able to complete the remaining work within the 70 hours budgeted. This will give you a ***favorable*** *variance of 10 hours* for the entire audit. There is a 50 percent probability of this outcome occurring.

Outcome 2. This quick pace of work will cause you to make errors. Since you must take time to correct these errors to be sure that the audit is done properly, this will cause you to spend a total of 130 hours in completing the audit. You will then finish the audit with an ***unfavorable*** *variance of 50 hours* for the total audit. There is a 50 percent probability of this outcome occurring.

Option B

Set the monthly budget at 100 hours. This is the amount of time originally budgeted to complete the remaining portion of the audit. If you select this monthly budget, you will be able to work at a normal work pace in completing the audit. This will mean that you have an ***unfavorable variance of 20 hours*** for the total audit. There is a 100 percent probability of this outcome occurring if you choose B.

I choose ________ (enter A or B).

Please indicate how much you prefer the choice you made in relation to the other one.

1-------2-------3-------4-------5
Slightly Moderately Very Much

References

Berg, J., L. Daley, J. Dickhaut, and J. O'Brien. 1986. Controlling preferences in lotteries on units of experimental exchange. *Quarterly Journal of Economics* 101 (May): 281–306.

Chang, O., D. Nichols, and J. Schultz. 1987. Taxpayer attitudes toward tax audit risk. *Journal of Economic Psychology* 8 (September): 299–309.

Cohen, M., J. Jaffray, and T. Said. 1987. Experimental comparison of individual behavior under risk and under uncertainty for gains and for losses. *Organizational Behavior and Human Performance* 39 (February): 1–22.

Deci, E. 1975. *Intrinsic Motivation.* New York: Plenum.

Fischhoff, B. 1982. Debiasing. In *Judgment Under Uncertainty: Heuristics and Biases,* edited by D. Kahneman, P. Slovic, and A. Tversky. New York: Cambridge Unversity Press.

Fishburn, P. 1977. Mean-risk analysis with risk associated with below-target returns. *American Economic Review* 67 (March): 116–26.

Hershey, J., and P. Schoemaker. 1980. Prospect theory's reflection hypothesis: A critical examination. *Organizational Behavior and Human Performance* 25 (June): 395–418.

Hogarth, R., and H. Einhorn. 1990. Venture theory: A model of decision weights. *Management Science* 36 (July): 780–803.

Kahneman, D., and A. Tversky. 1979. Prospect theory: An analysis of decision under risk. *Econometrica* 47 (March): 263–91.

Kaplan, R. 1982. *Advanced Management Accounting.* Englewood, NJ: Prentice-Hall, Inc.

Lakonishok, J., and S. Smidt. 1986. Volume for winners and losers: Taxation and other motives for stock trading. *The Journal of Finance* 41 (September): 951–74.

Leopard, A. 1978. Risk preference in consecutive gambling. *Journal of Experimental Psychology: Human Perception and Performance* 4 (August): 521–28.

Levin, I., R. Johnson, and M. Davis. 1987. How information frame influences risky decisions: Between-subjects and within-subject comparisons. *Journal of Economic Psychology* 8 (March): 43–54.

Libby, R., and P. Fishburn. 1977. Behavioral models of risk taking in business decisions: A survey and evaluation. *Journal of Accounting Research* 15 (Spring): 272–92.

Locke, E. A., and D. M. Schweiger. 1979. Participation and decision making: One more look. *Research and Organizational Behavior* Vol. 1: 265–339.

Lowenstein, G. 1988. Frames of mind in intertemporal choice. *Management Science* 34 (February): 200–14.

——, and D. Prelec. 1991. Negative time preference. *American Economic Review* 81 (May): 347–52.

Lopes, L. 1984. Risk and distributional inequality. *Journal of Experimental Psychology: Human Perception and Performance* 10 (August): 465–85.

Metzger, M. 1985. Biases in betting: An application of laboratory findings. *Psychological Reports* 56 (June): 883–88.

Payne, J., D. Laughhunn, and R. Crum. 1980. Translation of gambles and aspiration level effects in risky choice behavior. *Management Science* 26 (October): 1039–60.

——, ——, and ——. 1981. Further tests of aspiration level effects in risky choice behavior. *Management Science* 27 (August): 953–58.

Petty, J., and D. Scott. 1981. Capital budgeting practices in large American firms: A retrospective analysis and update. In *Financing Issues in Corporate Project Selection*, edited by F. Derkinderen and R. Crum. Boston: Martinus Nijhoff.

Rachlin, H. 1990. Why do people gamble and keep gambling despite heavy losses? *Psychological Science* 1 (September): 294–97.

Schadewald, M. 1989. Reference point effects in taxpayer decision making. *The Journal of the American Taxation Association* 10 (Spring): 68–84.

Schepanski, A., and D. Kelsey. 1990. Testing for framing effects in taxpayer compliance decisions. *The Journal of the American Taxation Association* 12 (Fall): 60–77.

Schneider, S., and L. Lopes. 1986. Reflection in preferences under risk: Who and when may suggest why. *Journal of Experimental Psychology: Human Perception and Performance* 12 (November): 535–48.

Schoemaker, P. 1982. The expected utility model: Its variants, purposes, evidence and limitations. *Journal of Economic Literature* 20 (June): 529–63.

———. 1990. Are risk-attitudes related across domains and response modes? *Management Science* 36 December): 1451 –63.

Seligman, M. 1975. *Helplessness: On depression, development and death.* San Francisco, CA: W. H. Freeman.

Shefrin, H., and M. Statman. 1985. The disposition to sell winners too early and ride losers too long: Theory and evidence. *The Journal of Finance* 40 (July): 777–90.

Tversky, A., and D. Kahneman. 1981. The framing of decisions and the psychology of choice. *Science* 211 (January): 453–58.

Thaler, R., and E. Johnson. 1990. Gambling with the house money and trying to break even: The effect of prior outcomes on risky choice. *Management Science* 36 (June): 643–60.

Waller, W. 1988. Slack in participative budgeting: The joint effect of a truth-inducing pay scheme and risk preferences. *Accounting, Organizations and Society* 13 (February): 87–98.

Weitzman, M. 1976. The new soviet incentive model. *Bell Journal of Economics* 7 (Spring): 251–57.

Young, M. 1985. Participative budgeting: The effects of risk aversion and asymmetric information on budgetary slack. *Journal of Accounting Research* 23 (Autumn): 829–42.

A FORUM ON
LIFO CHOICE OF INVENTORY VALUATION

1. Sasson Bar-Yosef and Pradyot K. Sen
   On Optimal Choice of Inventory Accounting Method

2. Ross Jennings, David P. Mest, and Robert B. Thompson, II
   Investor Reaction to Disclosures of 1974–75 LIFO Adoption Decision

3. Barry E. Cushing and Marc J. LeClere
   Evidence on the Determinants of Inventory Accounting Policy Choice

---

Editor's Comment: A "Forum" is a collection of articles around a common theme. Accepted refereed articles are generally followed by an invited "perspective" that offers a synthesis and suggestions for future directions of research on issues related to that theme. I had invited a group of professors to write two perspectives on "LIFO choice" but, for various reasons, closure could not be achieved. Thus, this "Forum" stands incomplete.

I had hoped the "perspective" would examine the extent to which the choice of LIFO or alternative inventory valuation methods is viewed as a signal of management's quality and strategy. For example, in what way does that choice describe a "corporate personality"? How does such a choice fit in the firm's portfolio of disclosure and incentive mechanisms? The articles included in this forum address some of these issues and advance our understanding of them a bit, but they do not help in offering full appreciation of the LIFO puzzle. We continue to be relatively uninformed about these issues and know little about the real reasons that many firms do not switch to LIFO when it appears that they would benefit by positive tax savings. Although the uncertainty about the LIFO puzzle is not reduced significantly, the three articles included in this Forum raise some interesting questions and offer insights for future research.

THE ACCOUNTING REVIEW
Vol. 67, No. 2
April 1992
pp. 320–336

On Optimal Choice of Inventory Accounting Method

Sasson Bar-Yosef*
The Hebrew University of Jerusalem
Pradyot K. Sen**
The University of California, Berkeley

SYNOPSIS AND INTRODUCTION: Inventory accounting rules may have a purchase (production) repercussion on the firm that affects stockholders' wealth, because managers, paid by an accounting income-based contract, can potentially distort the operation while maximizing their payoffs. Stockholders, by directly or indirectly dictating the accounting policy choice, can use accounting rules to implement their preferences and control the managers.

Analysis of such a scenario indicates that the stockholders' preferences towards accounting rules are driven primarily by three factors, the tax rate, the acquisition price, and the rate of real holding gain. With no taxes, and absent inflation, FIFO (first in, first out) implements the first-best decision rule. When taxes are present, the preferences of owners and managers diverge, which may lead to operating distortions.

An optimal accounting policy choice decision must consider this distortion (the incentive effect) as well as the tax gain possible with LIFO (last in, first out). Further analysis shows that FIFO may still be preferred over LIFO under certain conditions because an incentive effect may exceed an "obvious" tax advantage with LIFO. Abdel-khalik (1985) has provided empirical evidence suggesting that there are incentives for corporate executives to stay on FIFO despite the apparent LIFO tax advantages possibly because their incentives are income-based. Larcker (1983) provides independent evidence that often management compensation contracts are indeed based on accounting income. Thus, when incentive considerations are present, the cash flow effects as shown by Sunder (1976) may not be

---

* Partial financial support from the Kruger Center is gratefully acknowledged.
** Partial financial support from the Professional Accounting Program is gratefully acknowledged.

We thank Jeff Callen, Stefan Reichelstein, I. Venezia, three anonymous referees, the associate editors, and the editor of the *Review* for their useful comments.

Submitted February 1991.
Accepted November 1991.

obvious, and our analysis questions the conventional wisdom that firms tend to switch to LIFO at the time of rising prices.

Although the effect of LIFO versus FIFO accounting rule on firm cash flow is well-known,[1] many firms continue to follow FIFO in spite of tax advantages with LIFO.[2] Granof and Short (1984) reported that more than 30 percent of firms in their non-LIFO sample rejected LIFO because of "excessive cost" or "other adverse consequences." We argue that investors choose an inventory accounting method to optimize the trade-off between incentive effect and tax gain, and align the managers' interests as much as possible with those of the stockholders.[3]

Key Words: ***LIFO/FIFO, Incentive effects, Accounting policy choice.***

WE first discuss the role of the manager in choosing accounting policy and argue that managers have very little real choice in this regard. Next, we describe the setting. In the third section, we construct the model and analyze the cost tradeoff between incentive effect and tax gain. Some concluding remarks follow.

I. Accounting Policy Choice and the Manager

How much real freedom do the managers have in making a self-serving accounting policy choice? This is an empirical issue. We assume that investors, anticipating the managers' motives, do not delegate the inventory accounting decision if such delegation leads to a dead-weight loss (from not using potential tax savings). The inventory accounting rule is observable, and a forcing contract should be possible, which would eliminate the stockholders' need to adjust (ex post) for uncontrollable inefficiency arising out of such choice.[4] Instead, given the limited access to information and the existence of the accounting system for other purposes, such as external reporting and taxes, stockholders can use the accounting rules as mechanisms to implement their objectives. Therefore, although it may appear that management is making the choice of accounting policy, they may be doing so *in concurrence with the stockholders*.[5] It is on the basis of this premise that we pursue our analysis.[6]

[1] When costs are rising and the firm pays corporate taxes, the firm's reported periodic income decreases (increases) when the firm adopts LIFO (FIFO) inventory valuation method. However, Sunder (1976) shows that the firm's cash flows increase (decrease) with LIFO (FIFO) under a very general set of conditions. One such condition is that the cost of capital be greater than the inflation rate.

[2] The empirical findings for LIFO adoption are quite clear: the greater the tax savings potential, the greater the probability that the firm will adopt the LIFO inventory valuation method (Biddle 1980; Morse and Richardson 1983). Capital market reaction to a firm's decision to change their inventory valuation method from FIFO to LIFO is ambiguous; e.g., Abdel-khalik and McKeown (1978), Aharony and Bar-Yosef (1987), Biddle and Lindahl (1982), Brown (1980), Murray (1983), Ricks (1982), Stevenson (1987), and Sunder (1975) report conflicting evidence of LIFO adoption on stockholders' wealth.

[3] For a similar argument in the context of a choice of depreciation method see Ramakrishnan (1988).

[4] The above argument, of course, ignores the cost of writing such a forcing contract. But notice that the stockholders may not need an explicit contract to implement their preferred accounting method. They can voice their concerns during the firm's annual meeting and make their preferences known through voting.

[5] The manager must work according to the wishes of the majority of the stockholders, who have the power to remove the manager for nonperformance. Whether there is a conflict of interest between the majority and the minority stockholders is an interesting issue, but we do not consider it here.

[6] Delegating the choice of inventory accounting policy may allow managers to signal their private information (Bar-Yosef et al. 1991; Hughes and Schwartz 1988). We preclude that possibility in our model.

II. The Setting

Consider a setting involving resale of goods. Suppose the goods have uncertain demand distributed uniformly U (0, $\bar{x}$),[7] but it is common knowledge that there are two possible values of the maximum demand, $\bar{x}^H$ and $\bar{x}^L$, corresponding to two possible states of high and low demand. There are three periods in the model. During the initial period (0), the owner procures some goods (a quantity x_0 at a cost ϕ_0). At the beginning of the next period (1), the manager is hired to manage the operations. Upon hiring, the manager gets to know the demand state realization ($\bar{x}^H$ or $\bar{x}^L$). At the beginning of period 1, the manager purchases a quantity x_1 at a cost ϕ_1. In this period, the firm sells a quantity Q at a price p and incurs other operating expenses. We assume that the procurement and other operating costs are $c(x_1)$ and are incurred in cash. We assume that marginal costs are increasing at a non-decreasing rate, specifically, $c'(\bullet)>0$ and $c''(\bullet)\geq 0$.[8] Thus, if purchasing is delegated, the manager can arrive at a decision using more information. Without the manager, the owner must decide on the level of purchase on some average demand level based on prior information.[9] At the end of period 1, the manager leaves the employment and hands over excess inventory, if any.[10] The excess inventory is sold by the owner in period 2 for a price $k \bullet \phi_1$. Tax at the rate T is payable on current income at the end of period 1, and the tax on the profit on liquidation is paid at the end of period 2. In addition, we assume that both parties are risk-neutral.

At first, we assume that the manager is rewarded on the basis of income for the period.[11] Our main objective is to examine the impact of inventory accounting method on managerial decisions, given such an accounting income-based contract.[12] Subsequently, we consider a variation in the income-based contract where the manager can be made liable for the inventory loss incurred during the first period. Use of LIFO can potentially defer tax for one period when $\phi_1>\phi_0$. However, choice of an inventory accounting rule can alter costs and the performance measure, accounting income.

---

[7] We use uniform distribution for analytical tractability but our basic results can be interpreted equivalently for other probability distributions.

[8] The linear cost function $c(x)=a+b\bullet x$ satisfies the above conditions. Though not necessary, $c''(\bullet)\geq 0$ is sufficient for an interior solution of our problem.

[9] Depending on the manager's private information, it may not always be optimal to decide on the basis of additional information. Later, we establish the condition for the manager's information to have value, and restrict our analysis to that case.

[10] This assumption precludes creation of any long-term contract with the manager. The accounts with the manager must be settled when the relationship of the manager with the firm comes to an end. We consider this assumption realistic because the duration of contracts with the manager in real world is usually shorter than the life of the firm. Managers often leave the employment of a firm before the cash flow effects of their decisions are fully known, and their accounts are settled at the time they leave the employment.

[11] It is possible to construct a scenario with accounting income as the obvious candidate for writing contract. For example, if the owner observes only the assets at the beginning and at the end of the period (ending balances of cash and the inventory in this case), the difference between net assets (i.e., accounting income), must become the basis of all contracts. The manager's pay could be $w \bullet I$, where w is some weight and I, the accounting income. Such contracts could be invariant to the firm's accounting policy choice, as reported by Healy (1985). For some empirical evidence of such contracts, see Larcker (1983). Notice that an income-based contract captures the divergence of preference caused by the differences in time horizon of the owner and the manager.

[12] We do not explicitly model why the owner hires the manager and delegates operations, nor do we ask whether it is optimal to write a contract on income. Yet, notice that there is an economic demand for managers in our setting.

Therefore, as we show in the following model, an optimal rule must trade off tax savings and the cost associated with suboptimal decision making.

III. The Model

First, we establish when the manager's information has value. As described above, the manager starts with an initial inventory of x_0 units. Assessing the demand for output, the manager needs to decide the quantity of units to be purchased (denoted x_1). Let x_1^H and let x_1^L be the two optimal purchase policies corresponding to the high and low demand states that have maximum demand, $\bar{x}^H$ and $\bar{x}^L$, respectively. Let the cash flow in the high and low state be *CFH* and *CFL*, respectively, and *h*, the probability of the high state. If the owner could obtain the additional information, potentially the owner can generate an expected cash flow of $h \cdot CFH + (1-h)CFL$ and decide to purchase x_1^H with a probability *h* and x_1^L with a probability $(1-h)$. This is the benchmark (first-best) level of cash flow for the owner. Without additional information, let x_1^A be the owner's decision based on the expected demand (based on the prior distribution given by $\bar{x}^A = h\bar{x}^H + (1-h)\bar{x}^L$). Let the operating policy based on x_1^A lead to a cash flow denoted *CFA*. For the information to have value, it is necessary that:

$$h \cdot CFH + (1-h)CFL > CFA,$$

a condition we assume to exist. Even though the above condition is satisfied, it does not necessarily follow that the owner would continue to gain when the operating decision is delegated. If, however, the cash flow generated in the delegated regime approaches that of the left-hand side of the above inequality, the net gain to the owner becomes comparable to the first-best level of cash flow. To ensure a net gain to the owner in a delegated regime, it must be true that (the owner's share of) the weighted average of the cash flows associated with *maximized income* is greater than the average cash flow. Let *CFHIM* and *CFLIM* be the two cash flows in high and low demand states, respectively, when the manager maximizes income. A necessary and sufficient condition for the manager's information to have value in this setting is:

$$(1-m)[h \cdot CFHIM + (1-h) \cdot CFLIM] > CFA,$$

where *m* is the manager's share of the cash flow. Although this condition need not be satisfied in all cases,[13] we restrict our analysis to those cases where this condition holds. Intuitively, the above condition guarantees a demand for the manager and his or her private information.

We assume that it is possible to randomize between LIFO and FIFO, and denote δ as the fraction of inventory carried on FIFO basis. This assumption allows us to identify an optimal mixed strategy in the accounting policy choice. First, we show how the choice of inventory valuation rule may induce goal congruence in the absence of taxes. Since the accounting rule alters income and, thus, taxes, inducing the desired decision (LIFO) does not necessarily lead to a first-best solution. Therefore, we derive the second-best optimal choice of inventory accounting rule when taxes are present. Note that the computation method for inventory, income, and cash flow are identical for

---

[13] Notice that the above condition is a sufficient rather than a necessary condition for the manager's information to have value in this setting. For discussions of the necessary conditions for manager's information to have value, see Melumad and Reichelstein (1989).

Table 1
Calculation of Ending Inventory (*EINV*)

Demand (Q)	*Portion of FIFO*		*Portion of LIFO*
$0 \leq Q \leq x_0$	$\delta[(x_0-Q)\cdot\phi_0+x_1\cdot\phi_1]$	+	$(1-\delta)[x_0\cdot\phi_0+(x_1-Q)\cdot\phi_1]$
$x_0 \leq Q \leq x_1$	$\delta(x_0+x_1-Q)\phi_1$	+	$(1-\delta)[x_0\cdot\phi_0+(x_1-Q)\cdot\phi_1]$
$x_1 \leq Q \leq x_0+x_1$	$\delta(x_0+x_1-Q)\phi_1$	+	$(1-\delta)(x_0+x_1-Q)\phi_0$
$x_0+x_1 \leq Q < \bar{x}$	0		0
EINV	$\dfrac{x_0^2\cdot\phi_0+x_1^2\cdot\phi_1+2\phi_0\cdot x_0\cdot x_1}{2\bar{x}}$	+	$\delta(\phi_1-\phi_0)\dfrac{x_0\cdot x_1}{\bar{x}}$

Note: Expected cost of goods sold (*ECGS*)$=x_0\cdot\phi_0+x_1\cdot\phi_1-$*EINV*; Demand Q is distributed *U* $(0, \bar{x})$.

each demand state. Thus, for expositional ease, in this section we suppress the superscript as we describe the model. However, the demand states are used explicitly while maximizing the objective functions. Based on the acquisition policy and the respective period prices, the cost of goods available for sale (*CGAS*) are determined. These costs are expected to be allocated to the statement of income (as cost of goods sold, *CGS*) and to the statement of financial position (as cost of ending inventory, *INV*). The allocation depends on the inventory method assumed and the acquisition policy adopted by the manager. Though the manager gets to know which demand regime is operating, the actual demand and actual number of units of inventory are known only at the end of period 1. Let us consider the inventory costs in four mutually exclusive states of nature, which depend on the acquisition policy and the actual demand, as shown in table 1.[14] Here, the relationships among the actual demand Q, beginning inventory x_0, units available for sale (x_0+x_1), and the allocation of *CGAS* between *CGS* and *INV* are determined for these four distinctly different partitions of each state of nature.

On the basis of these four possible outcomes (of realization of Q), the acquisition decision (of x_1 units), and the accounting rule, the owner, upon hiring the manager, can compute the expected cost of goods sold (*ECGS*) and the expected ending inventory (*EINV*), as set out in table 1. Thus, if the manager adopts δ proportion FIFO and the actual demand for the output falls short of the number of units acquired (the second state), then *CGS* equals:

$$\delta[x_0\phi_0+(Q-x_0)\phi_1]+(1-\delta)Q\phi_1,$$

and the ending inventory becomes:

$$\delta(x_1+x_0-Q)\phi_1+(1-\delta)[(x_1-Q)\phi_1+x_0\phi_0].$$

Given these states of nature, the value of *EINV* may be computed as reported in table 1. Given that scenario, an owner, who could act directly, would maximize the expected end-of-period cash flow (*ECF*), which could be written as:

---

[14] Notice that though we suppress the H and L superscript denoting a high or a low demand state, our analysis holds independently for the two different states; i.e., there are four realizations each for two different demand regimes.

$$\max_{x_1}\int_0^{x_0+x_1} p\cdot Qf(Q)dQ - x_1\cdot\phi_1 - c(x_1) + [1-F(x_0+x_1)]\cdot p\cdot(x_0+x_1)$$

$$-T\left[\int_0^{x_0+x_1} p\cdot Qf(Q)dQ + [1-F(x_0+x_1)]\cdot p\cdot(x_0+x_1) - ECGS - c(x_1)\right] + LQ, \quad (1)$$

where $f(Q)$ and $F(Q)$ denote the density and the distribution functions of Q, respectively, and LQ is the expected cash flow from liquidating the ending inventory, given by:

$$LQ = \alpha(MINV - T[MINV - EINV]),$$

where $MINV$ is the market value of the ending inventory in units ($UINV$) given by $MINV = k\cdot\phi_1\cdot(UINV)$, k is a liquidation price factor, $k>0$, and α is the owner's present value discount factor. For the first-best level of cash flow, the maximization problem will be solved for the high demand state with a probability h and will be solved for the low demand state with a probability $1-h$.

The Income-based Contract

At first we consider only an income-based contract for the manager. The coefficient of T within the square bracket in equation (1) is the pre-tax expected net income (ENI) on which the income tax at the rate T is assessed. That is also the manager's objective function (i.e., the current period income), according to this compensation scheme. The manager, therefore, chooses the acquisition level, x_1, that will maximize this objective, given by:

$$\max_{x_1}\int_0^{x_0+x_1} p\cdot Q\cdot f(Q)dQ + [1-F(x_0+x_1)]\cdot p\cdot(x_0+x_1) - ECGS - c(x_1), \quad (2)$$

for each value of the demand state high and low. Noting that $CGAS$ is allocated to $ECGS$ and $EINV$, that is, $ECGS + EINV = x_0\cdot\phi_0 + x_1\cdot\phi_1$, and $f(Q)$ is distributed as $U(0,\bar{x})$, we can simplify ECF and ENI as follows:

$$G = ECF = \left[\left\{(x_0+x_1) - \frac{(x_0+x_1)^2}{2\bar{x}}\right\}\cdot p + \frac{(x_0+x_1)^2}{2\bar{x}}k\alpha\phi_1 - c(x_1)\right](1-T)$$

$$-x_1\phi_1 + T(x_0\cdot\phi_0 + x_1\cdot\phi_1) - T(1-\alpha)EINV, \quad (3)$$

and

$$M = ENI = \left[(x_0+x_1) - \frac{(x_0+x_1)^2}{2\bar{x}}\right]\cdot p - (x_0\cdot\phi_0 + x_1\cdot\phi_1) - c(x_1) + EINV. \quad (4)$$

The conflict of interest between the owner and the manager is apparent from the above objective functions; the owner's objective function ECF *is decreasing in* $EINV$ while the manager's objective function ENI is increasing in $EINV$.[15] The impact of the procurement policy on both the manager and the owner can be examined through the first-order conditions of equations (3) and (4). *Ex ante*, the owner expects the manager to receive a news of high demand with a probability h and a news of low demand with a probability $1-h$. Therefore, the owner expects the manager to solve the above maximization problem for the high (low) demand state with a probability h (correspondingly $1-h$), generating a net income of *ENIH* (*ENIL*). The first-order conditions for maximizing *ECF* under each demand state as a function of the accounting rules is given by:

---

[15] Though apparently similar, our characterization of preferences based on accounting and cash flow are different from that in Amershi and Sunder (1987), where the manager's belief about the owner's decision rule created the disparity. We have no such divergence of beliefs.

$$\left(1-\frac{x_0+x_1}{\bar{x}}\right)\cdot p+\phi_1\left[\frac{x_0+x_1}{\bar{x}}k\alpha-\frac{T(1-\alpha)(\delta x_0+x_1)}{(1-T)\bar{x}}\right]$$
$$-\frac{Tx_0\phi_0}{(1-T)\bar{x}}(1-\alpha)(1-\delta)=\phi_1+c'(x_1), \tag{5}$$

and the first-order condition for maximizing the *ENI* is given by:

$$\left(1-\frac{x_0+x_1}{\bar{x}}\right)\cdot p+\phi_1\left(\frac{x_0+x_1}{\bar{x}}\right)-(1-\delta)\frac{x_0(\phi_1-\phi_0)}{\bar{x}}=\phi_1+c'(x_1). \tag{6}$$

The interpretation of the above expressions is straightforward. In each, the marginal cost (the left-hand side of each equation) is being equated with the marginal benefits (the right-hand side of each equation). For example, in the absence of taxes (i.e., setting $T=0$), the owner equates the marginal cash outflow, $\phi_1+c'(x_1)$, to the marginal cash inflow, which is,

$$\{\text{Prob. }(\text{demand}\geq x_0+x_1)\}\cdot p+\{\text{Prob. }(\text{demand}<x_0+x_1)\}k\cdot\alpha\cdot\phi_1.$$

The manager's condition also reflects a similar calculation of cost and benefit.

Contract Based on Income and Liquidation Loss

It is conceivable that the owner would hold the manager responsible for any inventory loss at the last period, and therefore, the manager's contract can include an ex post settlement clause to adjust for losses caused by any inventory liquidation. Specifically, the manager's objective function in equation (2) could be modified to include a negative term equal to the inventory liquidation loss (INVLOSS) as follows:

$$\max_{x_1}\int_0^{x_0+x_1} p\cdot Q\cdot f(Q)dQ+[1-F(x_0+x_1)]\cdot p\cdot(x_0+x_1)-ECGS-c(x_1)$$
$$-\phi_1\frac{(x_0+x_1)^2}{2\bar{x}}(1-k\alpha). \tag{7}$$

Equation (7) can be simplified to a form similar to that of equation (4) with the additional term at its end, which would be of the form *ENI*−*INVLOSS*.

Inventory Accounting Policy without the Tax Effect

First, note that the first-order condition of the owner in equation (5) becomes independent of δ if there are no taxes (T becomes equal to zero), which makes the owner indifferent toward accounting policy. Further, if $k\alpha=1$ or, equivalently, with no real discounting, the two conditions (5) and (6) become identical. It is so because in a one-period economy, since terminal cash flow can be observed (and contracted upon), accounting has no role to play. In terms of delegated decision, the owner prefers an accounting rule that generates the most desirable cash flow. Without taxes, it is possible for the owner to induce a first-best allocation through the choice of accounting rule, as noted in the following proposition.

Proposition 1: With an income-based contract, the following choice of inventory accounting policy by the owner leads to a first-best decision by the manager.

$$\delta^* = 1 - \frac{\phi_1\{hx^H + (1-h)x^L\}}{x_0(\phi_1-\phi_0)\left(\frac{h}{\bar{x}^H} + \frac{(1-h)}{\bar{x}^L}\right)}(1-k\alpha), \tag{8}$$

where:

$$x^j(j=H,L) = \frac{(x_0+x_1{}^j)}{\bar{x}^j}. \tag{9}$$

With an income-based contract that is adjusted for liquidation loss, FIFO is always optimal and leads to a first-best decision.

Proof: Maximize $h \bullet CFH + (1-h) \bullet CFL$ for the owner and substitute $T=0$ from equation (5) to find the owner's first-order condition for maximization:

$$p[1-\{hx^H+(1-h)x^L\}] + k\alpha\phi_1\{hx^H+(1-h)x^L\} = \phi_1 + \{hc'(x_1^H)+(1-h)c'(x_1^L)\}. \tag{10}$$

With only an income-based contract, the owner expects the manager to maximize $h \bullet ENIH + (1-h)ENIL$. Use equation (6) to find the corresponding first-order condition for the manager;

$$p[1-\{hx^H+(1-h)x^L\}] + \phi_1\{hx^H+(1-h)x^L\} - (1-\delta)\{x_0(\phi_1-\phi_0)\}\left\{\frac{h}{\bar{x}^H}+\frac{1-h}{\bar{x}^L}\right\} = \phi_1 + \{hc'(x_1^H)+(1-h)c'(x_1^L)\}. \tag{11}$$

Comparison of equations (10) and (11) establishes that the right-hand sides of equations (10) and (11) are identical, and if:

$$k\alpha\phi_1\{hx^H+(1-h)x^L\} = \phi_1\{hx^H+(1-h)x^L\} - (1-\delta)\{x_0(\phi_1-\phi_0)\}\left\{\frac{h}{\bar{x}^H}+\frac{1-h}{\bar{x}^L}\right\}, \tag{12}$$

then the left-hand sides are also identical, which leads to goal congruence. Rearrangement of terms gives the required condition.

For adjustment of liquidation loss, deduct:

$$\phi_1\left[\frac{(x_0+x_1)^2}{2\bar{x}}\right](1-k\alpha),$$

from the manager's objective function and notice that the first-order condition with FIFO is exactly identical with that of the owner.

Two extreme cases of the above proposition give us conditions when either FIFO or LIFO may lead to goal congruence with the income-based contract.

Corollary 1: With an income-based contract,

(a) If $k \bullet \alpha = 1$, then choose FIFO,

(b) If $k \bullet \alpha = 1 - \dfrac{x_0(\phi_1 - \phi_0)}{\phi_1[hx^H + (1-h)x^L]}\left(\dfrac{h}{\bar{x}^H} + \dfrac{(1-h)}{\bar{x}^L}\right)$, then choose LIFO,

with x^j as defined in Proposition 1 before.

Proof: Check that the specified values, when substituted in equation (8) leads to $\delta = 1$ and 0, respectively.

Two observations are in order. First, note that the owner reaches the first-best solution by allowing the manager to pursue his or her own utility maximization. With an inventory loss adjustment clause, the owner can always get a first-best decision with FIFO. With a contract written on income alone, for some parameter values, such first-best allocation is also possible. In either case, the owner need not know the demand state faced by the firm. This result (with the contract with inventory loss adjustment) is rather strong because any purchase decision with prior information by the owner would result in a dead-weight loss. The same is also true for the income-based contract whenever $k\alpha = 1$. *Thus, despite an informational disadvantage, by delegating the purchase decision to the manager but controlling the inventory accounting rule, the owner can benefit from manager's superior information in these regimes.* In the case of an income-based contract, the goal congruence leads to an *ex ante* first-best solution in the sense that the owner can do no better alone. There will, however, be some *ex post* dead-weight loss in these cases whenever FIFO is not optimal.

Second, the above proposition highlights the role of the estimated value of the discounted disposal price of the excess inventory (in the following period) in inventory accounting policy choice. The discounted real disposal price ($k\alpha\phi_1$) decides the extent to which goal congruence is possible through inventory accounting policy choice. The factor $k \bullet \alpha$ is also the real holding gain factor. This factor determines the speculative demand of an item that the owner may want to consider while planning the inventory. Notice that in a period of stationary real prices, $k\alpha\phi_1$ is equal to ϕ_1, which makes $k\alpha = 1$. Therefore, if $k\alpha$ is equal to 1, that signifies absence of any speculative demand for the item. If so, as our result indicates, FIFO accounting rule leads to a first-best solution whenever income-based incentive contracts are used. If, however, parameter values lead δ^* to be ≥ 1 (or ≤ 0), a decision of FIFO (or LIFO) leads to a second-best solution. Notice that, under conditions of Proposition 1, markdown to market cannot occur.[16]

---

[16] Since α is less than 1, $k\alpha = 1$ signifies that k is greater than 1, which implies that the nominal disposal prices are greater than the purchase price. Although less likely, in times of rising prices that is possible. But, for LIFO to be optimal, k must be less than 1 according to our condition (b) of Corollary 1. Note that k less than 1 does not necessarily imply that inventory needs to be adjusted to the market value. That is because the cost of ending inventory is a weighted average of the beginning inventory cost and the purchase price. Thus, with a FIFO rule, when the demand Q is greater than the initial inventory x_0 (i.e., a penetration in the new layer occurs), $k < 1$ would imply a markdown to market. With a LIFO rule, markdown to market occurs if $k\phi_1$ is less than some weighted average of ϕ_0 and ϕ_1, depending on the level of ending inventory. Notice that with k less than 1, according to Corollary 1(a), FIFO will no longer be an optimal choice. Further, Corollary 1(b) implies

The optimal accounting policy δ^* is increasing in the parameter values of k, α, and ϕ_1, as stated in the following Proposition 2.

Proposition 2: The optimal inventory accounting policy δ^* of Proposition 1 is increasing in k, α, and ϕ_1.

Proof: Differentiate equation (8) with respect to k, α, and ϕ_1 and note that the signs of the respective derivatives of δ^* are all positive.

It follows from the above that the incentive effects make FIFO (LIFO) optimal at the time of *rising* (*falling*) prices, which is exactly opposite to the tax effect. It is, therefore, conceivable that when the tax effects are "small" in some sense, the decision of inventory accounting policy could be dominated by the incentive effects.[17] In the next section, we characterize the optimal inventory accounting policy choice when taxes are present.

Inventory Accounting Policy with the Tax Effect

When taxes are present, different accounting rules place the owner in a different cash flow regime. Therefore, though it is possible, implementing the first-best rule may not necessarily lead to a first-best level of cash flow. Thus, one needs to compare the levels of cash flow across different regimes when accounting policies are altered. In this section, we focus our analysis on the choice between LIFO and FIFO. Further, we shall assume that $\phi_1 > \phi_0$ so that LIFO tax advantages exist. Let the value of $k\alpha\phi_1$ be such that the condition for Proposition 1 holds so that the optimal accounting policy choice is between FIFO and LIFO. For ease of exposition, we first present a simple case. If there were no demand state uncertainty (i.e., if there were no information asymmetry), the condition under which the owner prefers FIFO to LIFO even with tax advantages is given in Proposition 3.

Proposition 3: With no demand state uncertainty, the necessary and sufficient condition for the owner to prefer FIFO over LIFO is:

$$\left[p\left(1-\frac{x_0+x_{1fl}}{\bar{x}}\right)-\left(\frac{1-T\alpha}{1-T}-k\alpha\right)\frac{x_0+x_{1fl}}{\bar{x}}\bullet\phi_1\right](x_{1f}-x_{1l})-\Delta c(x_{1f},x_{1l})$$

$$\geq\frac{T}{1-T}(1-\alpha)\frac{x_0\bullet x_{1l}}{\bar{x}}(\phi_1-\phi_0), \tag{13}$$

where:

$$x_{1fl}=\frac{x_{1f}+x_{1l}}{2}$$

---

that LIFO is optimal if $k\alpha\phi_1$ is minimum within the admissible range and is equal to some overall weighted average cost of the goods available for sale. For any intermediate value of $k\alpha\phi_1$, a mixed accounting policy of FIFO and LIFO is optimal, as indicated in proposition 1. Though it is possible to construct an incentive-compatible inventory accounting rule based on market valuation, we do not attempt to derive it here.

[17] In their analysis, Hughes and Schwartz (1988) ignore the costs associated with the incentive effect. To the extent these costs are large, the incentives to capture the tax differential, and hence to signal bad news through LIFO adoption, will become lesser.

and x_{1f}, x_{1l} are the optimum quantities purchased by the manager under FIFO and LIFO, respectively, and $\Delta c(x_{1f},x_{1l})$ = incremental procurement (production) cost for x_{1f} over x_{1l} as defined before.

Proof: The expected cash flow effect due to taxes when changing inventory accounting policy is:

$$\frac{T}{1-T}(1-\alpha)\Delta EINV,$$

where Δ *EINV* is the incremental inventory associated with the change in accounting policy. Substituting the values of *EINV* for FIFO and LIFO yields the right-hand side of equation (13). The left hand side of equation (13) is the gain associated with a better decision, which is calculated as:

$$\int_{x_{1f}}^{x_{1l}} \frac{\partial ECF(FIFO)}{\partial x_1} dx_1$$

and can be computed from the owner's first-order condition, given *FIFO*. Simplification establishes the inequality.

When demand state uncertainties are present and the manager has superior information, the owner can design the accounting policy with a view to exploit it. The owner will allow for the fact that the manager will buy a quantity x_1^H or x_1^L that maximizes the manager's objective, given the demand information received. The objective and, consequently, the quantity purchased will be a function of the accounting policy chosen by the owner. Therefore, the owner will want to maximize his or her own utility (cash flow ECF) subject to the manager's utility maximization process. Assuming that the manager's compensation is based on net income (*ENI*) minus the loss on ending inventory (*INVLOSS*), the owner's problem can be formally stated as:

$$\text{Max} \quad G(\delta)=h \cdot ECF^H[x_1^H(\delta)]+(1-h) \cdot ECF^L[x_1^L(\delta)],$$

$$\text{Subject to: } x_1^H, x_1^L \in \text{argmax}[h \cdot (ENI^H-INVLOSS^H)+(1-h)(ENI^L-INVLOSS^L)].$$

We assume a linear cost function for this section. Specifically, we assume that our cost function is of the form $c(x)=a+b \cdot x$. The solution of the owner's optimization problem can be characterized by the following first-order condition:

$$\frac{\partial G(\delta)}{\partial \delta}=h \cdot \frac{\partial ECF^H}{\partial x_1^H} \cdot \frac{\partial x_1^H}{\partial \delta}+(1-h) \cdot \frac{\partial ECF^L}{\partial x_1^L} \cdot \frac{\partial x_1^L}{\partial \delta}. \tag{14}$$

By replacing the argmax condition with the respective first-order condition, the optimality condition of the manager's problem for each demand state can be written, suppressing the state superscript, as:

$$\frac{\partial M}{\partial x_1}=\left(1-\frac{x_0+x_1}{\bar{x}}\right)p+k\alpha\phi_1\frac{x_0+x_1}{\bar{x}}-(1-\delta)\frac{x_0}{\bar{x}}(\phi_1-\phi_0)-b-\phi_1=0. \tag{15}$$

Notice that the linear cost function assumption allows us to solve the optimal decision function of the manager from equation (15) as a linear function of the underlying information variables. Collect terms to write the optimal decision $x_1(d;\delta)$ as:

$$x_1(d;\delta)=\beta_0+\beta_1 d+\beta_2\delta \tag{16}$$

where:

$$\beta_0 = -x_0\frac{p-\phi_0}{p-\phi_1},\ \beta_1 = \frac{p-\phi_1-b}{p-k\alpha\phi_1},\ \beta_2 = \frac{\phi_1-\phi_0}{p-\phi_1}$$

and d is the information parameter that takes values $\bar{x}^H$ and $\bar{x}^L$. Notice further that the effects of inventory accounting rule and the demand parameters are independent in the above construction. We use these properties to solve for δ^* from the first-order condition of the optimization problem. The results are stated in Proposition 4.

Proposition 4: The optimal solution to the inventory accounting rule is given by:

$$\delta^* = \frac{[h\theta^H+(1-h)\theta^L](1-T)-T(1-\alpha)\{\phi_1[hx_1^H+(1-h)x_1^L]+x_0\phi_0\}}{Tx_0(1-\alpha)(\phi_1-\phi_0)}, \tag{17}$$

where:

$$\theta^j = (\bar{x}^j - x_0 - x_1^j)p + \phi_1[k\alpha(x_0+x_1^j)-\bar{x}^j] - b\bar{x}^j \qquad \text{for } j=H,L \tag{18}$$

and the purchase levels are the optimal purchase levels decided by the manager.

Proof: Use equation (3) to compute equation (14) as follows:

$$\begin{aligned}\frac{\partial G(\delta)}{\partial\delta} = h\Bigg\{&\left(1-\frac{x_0+x_1^H}{\bar{x}^H}\right)\bullet p+\phi_1\left[\frac{x_0+x_1^H}{\bar{x}^H}k\alpha-\frac{T(1-\alpha)(\delta x_0+x_1^H)}{(1-T)\bar{x}^H}\right]\\ &-\frac{Tx_0\phi_0}{(1-T)\bar{x}^H}(1-\alpha)(1-\delta)-\phi_1-b\Bigg\}\frac{\partial x_1^H}{\partial\delta}+(1-h)\Bigg\{\left(1-\frac{x_0+x_1^L}{\bar{x}^L}\right)\bullet p\\ &+\phi_1\left[\frac{x_0+x_1^L}{\bar{x}^L}k\alpha-\frac{T(1-\alpha)(\delta x_0+x_1^L)}{(1-T)\bar{x}^L}\right]\\ &-\frac{Tx_0\phi_0}{(1-T)\bar{x}^L}(1-\alpha)(1-\delta)-\phi_1-b\Bigg\}\frac{\partial x_1^L}{\partial\delta}=0\end{aligned} \tag{19}$$

Note that given linearity in equation (16),

$$\frac{\partial x_1^H}{\partial\delta} = \frac{\partial x_1^L}{\partial\delta} = \beta_2. \tag{20}$$

Substitute equation (20) in equation (19), define θ appropriately, cancel and collect terms to get the required expression.

Proposition 4 characterizes a second-best solution. In comparison with the no-tax case, it is clear that the main source of distortion is tax, and not the information asymmetry. Given taxes, a first-best solution is not attainable except by chance. Though the optimal solution is a mixed one, conditions exist when a pure strategy of either FIFO or LIFO could be optimal, as shown in the following Proposition 5.

Proposition 5: The necessary and sufficient condition for the owner to prefer FIFO is given by:

$$\frac{h\theta^H+(1-h)\theta^L}{\phi_1[hx_1^H+(1-h)x_1^L]+x_0\phi_1} \geq \frac{T}{1-T}(1-\alpha); \tag{21}$$

the owner prefers LIFO if:

$$\frac{h\theta^{H}+(1-h)\theta^{L}}{\phi_1[hx_1^{H}+(1-h)x_1^{L}]+x_0\phi_0}\leq\frac{T}{1-T}(1-\alpha), \tag{22}$$

where θ is defined as in Proposition 4.

Proof: For FIFO to be optimal δ^* must be ≥ 1. Set equation (17) to be greater than or equal to 1, and collect terms to get the result. Similarly, for LIFO to be optimal, δ^* must be ≤ 0. Set equation (17) to be less than or equal to zero and collect terms to get the result.

The idea of Proposition 5 can be better understood from figure 1.[16] Panels A and B in figure 1 illustrate different optimal accounting rule and show how the inventory accounting choice may affect the level of purchases and its subsequent effect on owner's wealth. The upper graphs in panels A and B illustrate the manager's optimal level of purchases. The decision is characterized by the intersection of marginal cost $\phi_1+c'(x_1)$ and the manager's marginal utility under either LIFO or FIFO. The lower graphs show the corresponding points of the owner's utility under LIFO or FIFO regime. If the accounting rule chosen is LIFO (FIFO) the manager purchases $x_{1l}(x_{1f})$ and the owner's utility is $l(f)$.

Although the managers' preferred levels of purchases increase monotonically over the accounting rule, the owner's preferred levels of cash flow over these accounting rules need not be monotonic over the levels of purchases or the accounting rules themselves. Depending on the shape of the cash flow function, as determined by the factor $k\alpha\phi_1$ and other parameters, any one of the two methods could be preferred by the owner, *given that the manager makes the delegated purchase decisions according to the manager's best interest.*

The owner's optimal inventory policy, δ^*, is a function of several economic parameters such as the tax rate (T), the acquisition cost (ϕ_1), the liquidation price factor (k), and the discount factor (α). In the following Proposition 6, we examine the effect of change in these parameters on the optimal inventory accounting policy.

Proposition 6: When an interior solution to the optimal inventory accounting policy exists such that $0\leq\delta^*\leq 1$, the optimal policy is affected by the economic parameters in the following way:

1. The lower the tax rate, the more the preference is for FIFO, i.e.,

$$\frac{\partial\delta^*}{\partial T}<0.$$

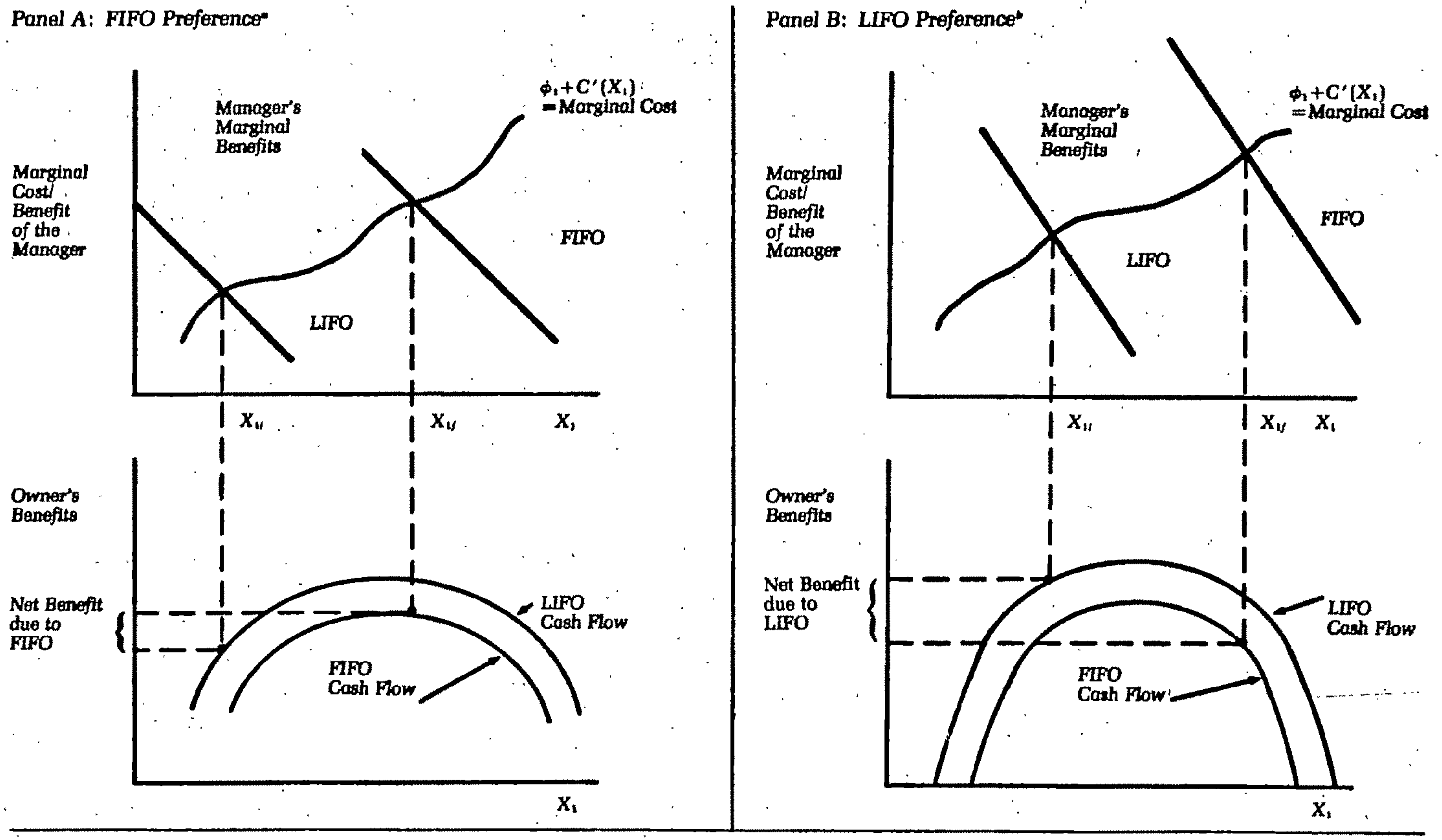
2. The higher the discount factor, the higher the preference is for FIFO, i.e.,

$$\frac{\partial\delta^*}{\partial\alpha}>0.$$

---

[16] Note that figure 1 represents a general cost function with $c'(x)>0$. If $c'(x)$ is a constant, as we assumed for Propositions 4 and 5, the upward sloping line representing $\phi_1+c'(x_1)$ will become a flat line.

Figure 1
Optimal Choice of FIFO and LIFO



Panel A: FIFO Preference[a]

Panel B: LIFO Preference[b]

---

[a] The net cash flow with FIFO is higher than that with LIFO although LIFO has a higher cash flow frontier. The incentive effect dominates the tax effect.

[b] The net cash flow with LIFO is higher than that with FIFO, as the higher cash flow (tax effect) dominates the incentive effect.

3. The higher the ending inventory liquidation factor, the higher the preference is for FIFO, i.e.,

$$\frac{\partial\delta^*}{\partial k} > 0.$$

4. The effect of the increase in the acquisition cost may result in a higher or a lower preference for FIFO, i.e.,

$$\frac{\partial\delta^*}{\partial\phi_1} > 0R < 0.$$

Proof: Tedious but straightforward differentiation of equation (17) establishes that the signs of the respective derivatives are unambiguous and the first three results follow. The derivative of δ^* with respect to ϕ_1 is given by:

$$\frac{\partial\delta^*}{\partial\phi_1} = \frac{k\alpha[x_0 + hx_1^H + (1-h)x_1^L](1-T) - T(1-\alpha)[hx_1^H + (1-h)x_1^L] - Tx_0(1-\alpha)\delta^*}{Tx_0(1-\alpha)(\phi_1 - \phi_0)}. \tag{23}$$

It is clear that conditions exist for equation (23) to be either greater or less than zero.

The above results are economically intuitive and are consistent with the no-tax case. An increase in tax rate shifts the trade-off towards the tax benefit, and the optimal inventory policy also shifts weight to LIFO. An increase in k or α reduces the likelihood of inventory loss in the following period and shifts the trade-off toward the incentive effect. The optimal inventory policy reflects that shift by putting more weight on FIFO. Finally, the effect of the acquisition cost is indeterminate in this setting because, as in the no-tax case, increase in acquisition cost increases the importance of the incentive effect but, with taxes present, the tax benefits also go up. Which of the two effects will dominate remains an empirical question.

IV. Discussions

We argue that when purchasing decisions are delegated, and managers are compensated by accounting income or some variation thereof, incentive problems could be minimized by appropriate choice of accounting policy. We show that the effect of accounting rule on the procurement decision need not be trivial and that sometimes the losses due to inefficient procurement may outweigh the gains on tax savings. By appropriately choosing the inventory accounting policy, the owner can achieve a goal congruence and, in some cases, such congruence can be achieved independently of the superior information of the manager. This is our explanation for why firms continue to follow FIFO in spite of tax advantages.

Our analysis implies that, given manager's incentives, inventory accounting method choice could lead to distortion of purchase and/or production decisions, thus altering the firm's cash flows. Such potential distortion could be the real cost of changing the inventory accounting method. Thus, the "switching cost" argument made in the literature can be viewed in this light as the incentive effects result from the accounting rules instead of the one-time implementation cost.[19]

Although we identify an optimal inventory accounting policy that calls for a mixed strategy in that the owner should use partly FIFO and partly LIFO, conditions exist for either FIFO or LIFO to be optimal. Though there are firms that use both FIFO and LIFO for parts of their inventory, the mixed policy could also be interpreted in other ways. One way would be to recognize that the weighted average accounting method would implement a particular mixed strategy of FIFO and LIFO. Yet another way the mixed strategy could be interpreted is as the propensity of a firm to use each of the two methods that captures the inherent probability of adopting each method. Although in real life we may observe either one or the other, a shift in this unobserved policy parameter δ may prompt a firm to switch from FIFO to LIFO.

Though most of our results agree with those established in the literature, one result is noteworthy in its difference. It is the effect of the acquisition cost on the firm's inventory accounting policy. A higher price, although it increases the tax advantages, also increases the value of potential distortion due to inefficient purchasing. Thus, contrary to existing theory, it is not clear from our analysis that, at the time of rising prices, firms will necessarily shift to LIFO.

[19] This is quite different from the "switching cost" argument offered in the empirical literature (e.g., see Biddle and Lindhal 1982) but could be a part of the "administrative cost" identified by Granof and Short (1984).

References

Abdel-khalik, A. 1985. The effect of LIFO-switching and firm ownership on executives' pay. *Journal of Accounting Research* 23 (Autumn): 427–47.

——, and J. C. McKeown. 1978. Understanding accounting changes in an efficient market: Evidence of differential reaction. *The Accounting Review* 53 (October): 851–68.

Aharony, J., and S. Bar-Yosef. 1987. Tests of the impact of LIFO adoption on stockholders: A stochastic dominance approach. *Contemporary Accounting Research* 3 (Spring): 430–44.

Amershi, A. H., and S. Sunder. 1987. Failure of stock prices to discipline managers in a rational expectations economy. *Journal of Accounting Research* 25 (Autumn): 177–96.

Bar-Yosef, S., P. J. Hughes, and I. Venezia. 1991. The LIFO/FIFO choice as a signal of future costs. University of California–Los Angeles, Mimeo.

Biddle, G. C. 1980. Accounting methods and management decisions: The case of inventory costing and inventory policy. *Journal of Accounting Research* 18 (Supplement): 235–80.

——, and F. W. Lindahl. 1982. Stock price reactions to LIFO adoptions: The association between excess returns and LIFO tax savings. *Journal of Accounting Research* 20 (Autumn): 551–88.

Brown, R. M. 1980. Short-range market reaction to changes to LIFO accounting using preliminary earnings announcement dates. *Journal of Accounting Research* 18 (Spring): 38–63.

Granof, M. H., and D. G. Short. 1984. Why do companies reject LIFO? *Journal of Accounting Auditing and Finance* 7 (Summer): 323–33.

Healy, P. M. 1985. The effect of bonus schemes on accounting decisions. *Journal of Accounting & Economics* 7 (April): 85–108.

Hughes, P. J., and E. S. Schwartz. 1988. The LIFO/FIFO choice: An asymmetric information approach. *Journal of Accounting Research* 26 (Supplement): 41–62.

Larcker, D. F. 1983. The association between performance plan adoption and corporate capital investment. *Journal of Accounting & Economics* 5 (April): 3–30.

Melumad, N., and S. Reichelstein. 1989. Value of communication in agencies. *Journal of Economic Theory* 47 (April): 334–68.

Morse, D., and G. Richardson. 1983. The LIFO/FIFO decision. *Journal of Accounting Research* 21 (Spring): 106–27.

Murray, D. 1983. The effect of certain research design choices on the assessment of the market's reaction to LIFO changes: A methodological study. *Journal of Accounting Research* 21 (Spring): 128–40.

Ramakrishnan, R. 1988. Accrual accounting and incentives: Depreciation methods for investment centers. Columbia University, Mimeo.

Ricks, W. E. 1982. The market's response to the 1974 LIFO adoptions. *Journal of Accounting Research* 20 (Autumn): 367–87.

Stevenson, F. L. 1987. New evidence on LIFO adoptions: The effects of more precise event dates. *Journal of Accounting Research* 25 (Autumn): 306–16.

Sunder, S. 1975. Stock price and risk related to accounting changes in inventory valuation. *The Accounting Review* 50 (April): 305–15.

———. 1976. Optimal choice between FIFO and LIFO. *Journal of Accounting Research* 14 (Summer): 323–33.

THE ACCOUNTING REVIEW
Vol. 67, No. 2
April 1992
pp. 337–354

Investor Reaction to Disclosures of 1974–75 LIFO Adoption Decisions

Ross Jennings
David P. Mest
University of Texas at Austin
Robert B. Thompson II
University of Maryland at College Park

SYNOPSIS AND INTRODUCTION: During 1974–75, more than a fourth of the manufacturing and merchandising firms listed on the New York Stock Exchange and the American Stock Exchange adopted or extended their use of the last-in, first-out (LIFO) method of inventory accounting. Because a change to LIFO accounting can materially affect a firm's cash flows, the stock price behavior of firms adopting LIFO during fiscal year 1974 has been studied extensively. However, existing studies provide little evidence that stock price movements during the period were directly linked with LIFO adoption.

In this article, we report the results of new tests for a stock price response to disclosures of 1974 LIFO decisions by both adopting and nonadopting firms. Our tests assume that, prior to disclosure, prices reflect available information about the likelihood of LIFO adoption. This implies (for both groups of firms) that the stock price change in response to disclosure of the LIFO decision depends on investors' revisions in their beliefs regarding the probability of adoption caused by the disclosure. We use this implication to design tests of the null hypothesis of no association between stock price movements and the disclosure of 1974 LIFO decisions.

Our tests are based on a sample of 487 firms that could have adopted LIFO during fiscal 1974. For each firm, we first identify an interval that includes the disclosure of its LIFO decision. We then estimate the probability of LIFO adoption at the beginning of that interval on the basis of infor-

We thank Dick Dietrich, Bill Kinney, Fred Lindahl, Frank Stevenson, Senyo Tse, and workshop participants at the University of California, Berkeley, George Mason University, the University of Maryland at College Park, and the University of Texas at Austin for helpful comments on previous drafts. We also thank the referees for their useful comments. We are grateful to Frank Stevenson for generously sharing his data.

Submitted May 1990.
Accepted July 1991.

mation that was available at the time. Finally, we regress disclosure-interval stock returns on revisions in the probability of adoption resulting from disclosure of the decision. Our results provide strong evidence of a price response for nonadopting firms and weaker evidence of a price response for adopting firms.

Key Words: *Accounting change, LIFO.*

Data Availability: *Data used in this study will be made available upon request.*

THE remainder of this article is organized as follows. Section I reviews previous studies of the stock price reponse to LIFO adoption disclosures. Section II presents a model of stock price change at the time a firm discloses the LIFO decision. In section III, we describe procedures used to select the sample of firms on which the study is based and to identify a disclosure interval for each firm. Section IV describes the procedure used to estimate the probability of adopting LIFO. Test results are reported in section V, and concluding remarks are offered in section VI.

I. Previous Research

Previous studies of stock price response to disclosures of 1974 LIFO decisions fall into two groups, depending on the type of research design employed.[1] The first group (Abdel-khalik and McKeown 1978; Brown 1980; Ricks 1982) examined mean returns for adopting firms over intervals that are likely to include the LIFO disclosure. The second group (Biddle and Lindahl 1982; Ricks 1986; Stevenson 1987) examined the association between disclosure-interval security returns and expected LIFO tax savings.

Studies in the first group compared excess returns for adopting firms over the month or week of presumed disclosure with the excess returns for nonadopting "control" firms during the same interval. Each of these studies reported a significant difference between the average excess returns for the two groups. However, neither group can realistically be viewed as a control for the other because firms "self-select" into adopting and nonadopting groups.[2] As evidence of this, differential return behavior was also observed for periods outside the disclosure interval in each of these studies. Thus, no inferences can be made about stock price response to LIFO disclosures from these studies.

Biddle and Lindahl (1982) reported a positive association between excess adoption-year stock returns and firm-specific estimates of LIFO tax savings, after controlling for unexpected earnings. They concluded that investors responded positively to the adoption of LIFO by pricing securities upward in proportion to the expected tax savings flowing from the accounting change. In a later study, however, Ricks (1986) demonstrated that the Biddle and Lindahl result may be due to a bias in measured abnormal returns that was correlated with LIFO tax savings.

---

[1] See Biddle and Lindahl (1982) and Lindahl et al. (1988) for reviews of these studies.

[2] This self-selection problem was first recognized by Sunder (1973), who studied the behavior of excess stock returns for a sample of pre-1974 LIFO adoptions.

Stevenson (1987), using a similar research design, examined excess returns over 2-day, 10-day, 60-day, and 250-day intervals that ended on a carefully researched disclosure date. After controlling for unexpected earnings and firm size, Stevenson found a significant association between returns and LIFO tax savings only for the two longer intervals. For these intervals, Stevenson measured unexpected earnings as the difference between reported LIFO earnings and a forecast of FIFO earnings. However, this measure of unexpected earnings understates unexpected LIFO earnings by an amount equal to the expected effect of LIFO adoption on earnings. Thus, Stevenson's tax savings variable may be significant simply because it is correlated with error in the measurement of unexpected earnings.[3]

II. Model and Research Design

In this study, we test the following hypothesis:

H_0: Investors did not respond to disclosures of fiscal 1974 LIFO decisions.

Our test assumes that investors are risk-neutral and that available information about the firm (including information bearing on the likelihood of LIFO adoption) is reflected in price.

Let V_{i0} be firm i's stock price just prior to disclosure of the LIFO decision. This price is a weighted average of the expected price if LIFO is adopted and the expected price if LIFO is not adopted:

$$V_{i0}=P_{i0}[E_0(V_i \mid \text{Adopt}, Z_0)]+(1-P_{i0})E_0(V_i \mid \text{Not Adopt}, Z_0), \tag{1}$$

where P_{i0} represents investors' assessment of the probability that the firm will adopt LIFO, and Z_0 is the set of publicly available information prior to disclosure.

If the *only* change in the set of publicly available data is the disclosure of the firm's LIFO decision, then the price will move to $V_{i1}=E_0(V_i \mid \text{Adopt}, Z_0)$ if the firm adopts, or to $V_{i1}=E_0(V_i \mid \text{Not Adopt}, Z_0)$ if the firm does not adopt. The resulting price change can be represented as:

$$\Delta V_i=B_i\Delta P_i, \tag{2}$$

where:

$$B_i=E_0(V_i \mid \text{Adopt}, Z_0)-E_0(V_i \mid \text{Not Adopt}, Z_0), \tag{3}$$

and ΔP_i is either $1-P_{i0}$ if the firm adopts LIFO, or $-P_{i0}$ if the firm does not adopt LIFO.

This model can be extended to allow for changes in firm value that are attributable to an "updated" information set (Z_1) as well as disclosure of the LIFO decision.[4] In this case, equation (2) becomes:

---

[3] To explore this issue further, we replicated Stevenson's analysis with our sample of adopting firms. In our replication, the dependent variable is size-adjusted returns measured over the disclosure intervals used in our study, which are generally longer than Stevenson's 10-day interval, but shorter than his 60-day interval. Like Stevenson, we find a significant positive association between excess returns and estimated LIFO tax savings when unexpected earnings is measured as "reported LIFO minus forecast FIFO" (t-statistic of 2.41). However, when unexpected earnings is measured on a consistent accounting basis, the estimated coefficient on expected tax savings reverses signs and is insignificant (t-statistic of -0.48).

[4] During 1974, "pure" LIFO disclosures were rare events. For example, in our sample, 81 percent of the LIFO disclosures accompanied an annual or quarterly earnings announcement. More generally, the longer the disclosure interval examined, the greater the likelihood that price will be influenced by information other than the LIFO disclosure.

$$\Delta V_i = A_i + B_i \Delta P_i, \tag{4}$$

where B_i is unchanged from equation (3), and A_i reflects the incremental value effect of new information other than the disclosure of the LIFO decision.[5]

To test the null hypothesis, we operationalize equation (4) by estimating the following cross-sectional regression:

$$\Delta V_i = \gamma_0 + \gamma_1 \Delta \hat{P}_i + \epsilon_i, \tag{5}$$

where $\hat{P}_i$ is an estimate of P_{i0}, and $\Delta \hat{P}_i$ equals $1 - \hat{P}_i$ for adopting firms and $-\hat{P}_i$ for nonadopting firms. In this regression, if $\Delta \hat{P}_i$ approximates investors' (unobservable) assessment of ΔP_i, then $\hat{\gamma}_0$ and $\hat{\gamma}_1$ estimate average values of A and B in equation (4) across firms in the sample. A finding that $\gamma_1 \neq 0$ serves as evidence to reject the null hypothesis.

A directional alternative to the null hypothesis cannot be constructed because, as discussed in Kang (1988) and Lanen and Thompson (1988), the price change at disclosure can be either positive or negative for both adopting and nonadopting firms. In appendix A, this argument is developed more formally with simple assumptions about the behavior of investors and managers.

III. Sample Selection

This study focuses on LIFO decisions for the fiscal year beginning June 1974 and ending May 1975. Each of the firms initially considered for inclusion in our final sample: (1) was listed on the New York Stock Exchange (NYSE) or the American Stock Exchange (AMEX) during the study period, (2) was covered in the first-quarter 1975 *Value Line Investment Survey*, and (3) reported inventories in fiscal year 1973. A group of 1,122 firms met these criteria (see table 1). From this group, we eliminated firms for which the LIFO decision was clearly unimportant. We also eliminated firms for which we were unable to obtain data needed to estimate the probability of LIFO adoption, including 266 firms for which we were unable to obtain price change data that were closely associated with the firm's four-digit SIC code. The final sample consists of 195 adopting firms (40%) and 292 nonadopting firms (60%). Each firm's classification was determined by examining its annual report or Form 10-K for fiscal year 1974. Both the industry composition and the proportion of adopting firms in the sample are similar to those reported by Ricks (1982).

Because it is difficult to identify LIFO disclosure dates precisely (Ricks 1982; Stevenson 1987), the shortest period over which the relevant price change can be observed is a "disclosure interval" that ends on the *earliest* identifiable disclosure date and begins on the *latest* date for which there is evidence that the firm's LIFO decision was not publicly known. To determine the ending date for this interval, we examined *The Wall Street Journal Index*, *The Wall Street Journal "Earnings Digest,"* and the *Value Line Investment Survey*.[6] If a public record of disclosure could not be located, the

---

[5] If the firm adopts LIFO, $A_i = E_1(V_i \mid \text{Adopt}, Z_1) - E_0(V_i \mid \text{Adopt}, Z_0)$. If the firm does not adopt LIFO, $A_i = E_1(V_i \mid \text{Not Adopt}, Z_1) - E_0(V_i \mid \text{Not Adopt}, Z_0)$. Although $B_i \Delta P_i$ represents the price response to disclosure of the LIFO decision *per se*, A_i may reflect other LIFO information, e.g., the LIFO earnings effect that becomes available during the disclosure interval.

[6] We were able to identify a number of disclosure intervals for adopting firms more precisely with data that were generously supplied by Frank Stevenson.

Table 1
Summary of Sample Selection Procedure

NYSE and AMEX firms covered by *Value Line* during first-quarter 1975, with inventories at 1973 fiscal year-end		1,122
Excluded:		
SIC codes 4000–4999, 6000–9999 (except 6711)	109	
Foreign firms	29	
Firms using LIFO as primary inventory method in 1973	92	
Firms using specific identification or current value as primary inventory method in 1973	18	
Firms not included in 1988 COMPUSTAT data base	53	
Data not available for 1969, 1970, 1971, or 1972	22	
Firms with change in fiscal year over the period 1969–74	27	
Firms without price change data	266	
Other*	19	635
Final sample		487

* "Other" includes firms with an annual earnings announcement date not listed in *The Wall Street Journal Index* (3), new listings in *Value Line* (4), delisted by exchange (1), previous agreement with IRS preventing adoption of LIFO (1), firm with a large merger during 1974 that resulted in a change in relative use of LIFO (1), firms that announced both a decision to adopt and a decision not to adopt LIFO during 1974 (4), and firms that were not traded on the day of their fiscal 1974 annual earnings announcement (5).

annual earnings announcement date was treated as the ending date of the disclosure interval.[7] After determining the ending date of the disclosure interval, we assigned the date of the immediately preceding *Value Line* report for the firm as the beginning date.[8] The mean (median) length of this interval for our sample is 29 (28) trading days.

---

[7] Because the Internal Revenue Code (section 472(c)) requires that firms adopting LIFO for tax purposes must also adopt LIFO for financial reporting purposes, the annual earnings announcement can be viewed as the last opportunity to disclose a LIFO decision for the fiscal year. The disclosure interval ends on an annual earnings announcement for 90 (285) adopting (nonadopting) firms. For the remainder of the adopting firms, the disclosure interval ends on a quarterly earnings announcement (21 firms), the publication date of a *Wall Street Journal* news article (42 firms), the publication date of a *Value Line* report (38 firms), or the publication date of a news article in some other publication (four firms). For the remaining seven nonadopting firms, the disclosure interval ends on the publication date of a *Value Line* report.

[8] During 1974–75, *Value Line* frequently discussed the possibility of LIFO adoption, and consistently reported the outcome of LIFO decisions disclosed since their last report. Thus, if *Value Line* had not reported a firm's LIFO decision, we assumed that the decision was not yet publicly known.

IV. Estimating the Probability of LIFO Adoption

To test the null hypothesis, we require an estimate of the probability of LIFO adoption for each sample firm, using information that was publicly available at the beginning of its disclosure interval.[9] To obtain these estimates, we followed the sequence of steps outlined below:

1. Determine for each firm in the sample the calendar week w_b in which its disclosure interval begins, and the calendar week w_e in which its disclosure interval ends.
2. For calendar week t, include all firms for which $w_e < t$ in the "week t estimation sample." Include all firms for which $w_b = t$ in the "week t disclosure sample."
3. Use firms in the week t estimation sample to estimate the logistic regression:

$$C_j = \mathbf{x}_j' b + \epsilon_j.$$

In this regression, C_j takes the value 1 if firm j is an adopting firm, and 0 otherwise; $\mathbf{x}_j$ is a vector of contemporaneously observable firm characteristics; $\boldsymbol{b}$ is a vector of regression parameters; and ϵ_j is an error term drawn from a logistic distribution.

4. Use the vector of parameter estimates $\hat{b}$ obtained from the week t estimation sample to predict:

$$\hat{P}_i = [1 + \exp(-\mathbf{x}_i' \hat{b})]^{-1}$$

for each firm in the week t disclosure sample. $\hat{P}_i$ is an estimate of the prior probability that firm i adopts LIFO, conditional on the information encoded in vector $\mathbf{x}_i$.

5. Repeat steps 2 through 4 for each week during the study period that contains the beginning of the disclosure interval for at least one sample firm.

Step 2 insures that $\hat{P}$ for a given firm is not influenced by information that became available after the beginning of its disclosure interval.[10] Each $\hat{P}$ produced by this procedure is an out-of-sample prediction that investors could have made at the beginning of the firm's disclosure interval. At the same time, because the estimation procedure "rolls forward" in calendar time, each $\hat{P}$ is based on the largest possible estimation sample.

Our estimation procedure cannot be applied to predict P until enough sample firms have disclosed LIFO decisions to estimate the model. By the week beginning August 19, 1974, 31 firms had disclosed LIFO decisions which allowed us to estimate P for the 416 firms whose return intervals had not yet begun.[11]

In implementing the procedure described above, we included in the vector $\mathbf{x}$ a large number of variables that are likely to be associated with the probability of LIFO adoption. For purposes of exposition, we divide these variables into four groups (see table 2). Variables in the first group are intended to capture the influence of inventory price changes on the probability of LIFO adoption. The second group contains measures of financial and operating characteristics that may be associated with the decision to adopt LIFO. Group 3 variables are based on evaluations contained in *Value Line Investment Survey* reports. The final group contains measures of the frequency of LIFO adoption in the firm's industry. A more detailed discussion of these variables is presented in appendix B. Table 3 reports mean and median values for each variable.

---

[9] A number of researchers have estimated probabilistic models of the LIFO decision (Cushing and LeClere 1992; Hunt 1985; Lee and Hsieh 1985; Lindahl 1989; Niehaus 1989) with results that suggest (but do not demonstrate) that anticipation of the outcome was possible.

[10] If we did not impose this restriction, we might obtain P estimates that are more "accurate" in terms of ex post classification. However, they would not necessarily be more accurate as a representation of investors' beliefs as of the beginning of any given firm's disclosure interval.

[11] This week falls in the middle of the disclosure interval for 40 firms. As a result, we are unable to estimate P for these firms.

Each of the variables described in table 2 was included in the vector **x** for purposes of estimating the logistic regression described in step 3 of our procedure. We conducted all analyses using two alternative specifications for the independent variables. In model A, continuous variables are not transformed or truncated in any way. In model B, continuous variables enter the model in ranked form. This assumes that only the order of the variables contains information relevant to the LIFO decision. Results are reported below for both models.

Our procedure required estimation of 44 weekly logistic regressions for both models, which resulted in 88 sets of coefficient estimates. Because the interpretation of individual regression coefficients is not relevant for this study, we do not report these estimates.[12]

Table 4 presents the distributions of $\hat{P}$ for adopting and nonadopting firms. Overall, the results indicate that nonadopting firms are more concentrated in the low $\hat{P}$ ranges, while adopting firms are more concentrated in the high $\hat{P}$ ranges. Also, for both models, the median $\hat{P}$ for adopting firms is about four times as large as that for nonadopting firms. Using a chi-square test for the difference in medians for two independent samples (Siegel and Castellan 1988, 124–28), we rejected the null hypothesis of no difference at the 0.001 level in both cases.

To assess the predictive power of the procedure, we predicted that firms with $\hat{P}>0.5$ would adopt LIFO and that those with a $\hat{P}\leq 0.5$ would not. Table 5 reports the frequency of decisions correctly predicted with this rule. For model A, 63 percent of the adopting firms and 70 percent of the nonadopting firms are correctly predicted. The results for model B are generally comparable. This evidence indicates that, prior to disclosure, investors could have used publicly available information to predict firms' 1974 LIFO decisions.

V. Investor Response to Disclosure of LIFO Decisions

Test Procedure

To test the null hypothesis, we estimated separate cross-sectional regressions analogous to equation (5) for nonadopting and adopting firms. In addition, we modified equation (5) to control for new earnings information released during the LIFO disclosure interval. Thus, for nonadopting firms, we estimated the following cross-sectional regression:

$$Ret_i=\gamma_0+\gamma_1\Delta\hat{P}_i+\gamma_2 UE_i+\epsilon_i, \qquad (6)$$

---

[12] In the final estimation of the logistic regression, coefficient estimates on the following variables were significant at or below the 10 percent level for both models: *HG73*, *INFLATION*, *VL1*, *VL2*, *CHANGE*, and *IND1*. In addition, coefficients on *INV*, *CGSSAL*, *CGSINV*, and *LTD* were significant for model A, and coefficients on *HG74*, *GFA*, *ROA*, *NOL*, and *IND2* were significant for model B.

Table 2
Predictor Variables Used to Estimate the Probability of LIFO Adoption

Group 1—*Direct measures of LIFO benefits*	
HG73	Estimate of the difference between LIFO and FIFO cost of sales for 1973 (using actual 1973 data) deflated by the market value of equity at the end of 1972.
HG74	Estimate of the difference between LIFO and FIFO cost of sales for 1974 (using data available prior to announcement of the LIFO decision) deflated by the market value of equity at the end of 1973.
INFLATION	Annualized rate of change in the four-digit industry price index for three months prior to announcement of the LIFO decision.
Group 2—*Financial and operating characteristics*	
INV	Ratio of inventory to assets, end of fiscal year 1973.
CGSSAL	Ratio of 1973 cost of goods sold to 1973 sales.
CGSINV	Ratio of 1973 cost of goods sold to end of fiscal year 1973 inventory.
CVINV	Coefficient of variation of inventory, 1969–73.
CLDBT	Ratio of current liabilities to total debt, end of fiscal year 1973.
FOREIGN	Dummy variable indicating that the firm's 1973 provision for income taxes included a foreign component.
ASSET	Total assets, end of fiscal year 1973.
GFA	Ratio of gross fixed assets, end of fiscal year 1973, to 1973 sales.
ROA	Return on 1973 assets (income before interest and taxes divided by year-end assets).
ROS	Ratio of 1973 total expense to 1973 sales.
CVINCOME	Coefficient of variation of pretax income, 1969–73.
NOL	Net operating loss carryforwards at the end of 1973 (per COMPUSTAT), deflated by 1973 sales.
LTD	Ratio of long-term debt to total assets, end of fiscal year 1973.
CR	Current ratio, end of fiscal year 1973.
OWNER	Percentage of common shares owned by managers and directors at the beginning of the disclosure interval.
Group 3—Value Line *evaluations*	
VL1	Dummy variable indicating that, according to *Value Line*, management was considering the adoption of LIFO.
VL2	Dummy variable indicating that, although *Value Line* did not report a considering statement by management, it did provide an analysis of the impact of a potential LIFO adoption.
CHANGE	Difference between the last *Value Line* forecast of 1974 earnings before the announcement of the LIFO decision and actual income for 1973, deflated by market value of equity at the end of 1973.
Group 4—*Frequency of LIFO adoption by two-digit industry*	
IND1	For each two-digit industry, the number of firms that used LIFO as their primary inventory method during fiscal year 1973 as a percentage of the sum of these firms and sample firms in the same industry.
IND2	Percentage of sample firms in two-digit industry that adopted LIFO during fiscal year 1974 prior to firm's LIFO decision announcement.

Table 3
Sample Means and Medians for Predictor Variables Used to Estimate the Probability of LIFO Adoptions

	Sample Mean			*Sample Median*		
*Predictor Variables**	*All Firms*	*Adopting Firms*	*Non-adopting Firms*	*All Firms*	*Adopting Firms*	*Non-adopting Firms*
HG73	0.04	0.04	0.05	0.02	0.02	0.01
HG74	0.20	0.25	0.17	0.10	0.13	0.07
INFLATION	0.27	0.38	0.19	0.17	0.27	0.14
INV	0.28	0.28	0.29	0.28	0.28	0.28
CGSSAL	0.70	0.72	0.68	0.72	0.73	0.71
CGSINV	5.31	4.42	5.90	3.74	3.90	3.58
CVINV	20.38	17.79	22.11	17.38	15.50	19.18
CLDBT	0.60	0.59	0.60	0.57	0.56	0.58
FOREIGN	0.52	0.52	0.51	1.00	1.00	1.00
ASSET	633.91	683.58	600.74	201.52	251.06	177.78
GFA	0.54	0.57	0.53	0.39	0.47	0.35
ROA	0.14	0.14	0.14	0.13	0.13	0.13
ROS	0.95	0.95	0.95	0.96	0.96	0.96
CVINCOME	91.37	53.05	116.96	29.70	30.18	28.63
NOL	0.01	0.00	0.01	0.00	0.00	0.00
LTD	0.19	0.18	0.19	0.19	0.19	0.19
CR	2.46	2.47	2.45	2.30	2.35	2.25
OWNER	0.14	0.12	0.16	0.05	0.03	0.08
VL1	0.12	0.27	0.02	0.00	0.00	0.00
VL2	0.07	0.11	0.04	0.00	0.00	0.00
CHANGE	0.04	0.09	0.01	0.01	0.02	0.01
IND1	0.10	0.13	0.07	0.07	0.08	0.06
IND2	0.12	0.11	0.13	0.08	0.08	0.08

* See table 2 for variable definitions.

where UE_i is a measure of the unexpected component of earnings information released during the interval. For adopting firms, we estimated:

$$Ret_i = \gamma_0 + \gamma_1 \Delta \hat{P}_i + \gamma_2 UE_i + \gamma_3 ULE_i + \epsilon_i, \tag{7}$$

where UE_i is a measure of unexpected earnings on a pre-change accounting basis, and ULE_i is the unexpected earnings effect of LIFO adoption.[13] In both regressions, *Ret* is measured as the difference between the disclosure interval return for each firm and the

---

[13] Estimating the regressions separately for adopting and nonadopting firms solves two potential econometric problems that could arise if a single equation were estimated. First, Ricks (1982) presents evidence that there were mean return differences between adopting and nonadopting firms at the fiscal 1974 earnings release date. In additon, $\Delta \hat{P}$ is always positive for adopting firms and always negative for nonadopting firms. Thus, if the regression is estimated for adopting and nonadopting firms together, these return differences may result in a significant coefficient on $\Delta \hat{P}$ that is unrelated to disclosure of LIFO decisions. Second, as discussed in section II, $\hat{\gamma}_1$ estimates the average value across the sample firms of the coefficient on $\Delta \hat{P}$. If this coefficient differs systematically between adopting and nonadopting firms, averaging across all firms may reduce the power of our test.

Table 4
Distribution of $\hat{P}$, the Estimated Probability of LIFO Adoption[1]

	Model A			Model B		
P-range	*Adopting Firms*	*Non-adopting Firms*	*Total*	*Adopting Firms*	*Non-adopting Firms*	*Total*
.00–.10	19.0%	35.7%	29.3%	13.3%	27.5%	22.1%
.10–.20	4.4	13.6	10.1	3.8	18.2	12.7
.20–.30	5.1	9.7	7.9	5.1	8.5	7.2
.30–.40	7.0	5.4	6.0	1.9	7.0	5.1
.40–.50	1.9	5.8	4.3	5.7	6.6	6.3
.50–.60	4.4	4.3	4.3	3.1	5.8	4.8
.60–.70	3.2	6.6	5.3	6.3	3.9	4.8
.70–.80	2.5	1.9	2.2	3.8	4.3	4.1
.80–.90	3.1	4.2	3.9	8.9	5.4	6.7
.90–1.00	49.4	12.8	26.7	48.1	12.8	26.2
	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%
Mean	0.639	0.338	0.453	0.693	0.371	0.493
First quartile	0.247	0.023	0.049	0.418	0.082	0.132
Median	0.886	0.211	0.346	0.882	0.232	0.432
Third quartile	1.000	0.609	0.954	1.000	0.640	0.929
Test for Difference in Medians[2]			38.0			54.4

[1] Statistics are based on the 416 sample observations for which P is estimated.

[2] The test statistic for difference in medians has a $\chi^2(1)$ distribution under the null hypothesis of no difference. The critical value for the 0.001 confidence level is 10.83.

Table 5
Prediction Results for $\hat{P}$

	Percent Correctly Predicted[1]				
Model	*Adopting Firms*	*Non-adopting Firms*	*Total*	*N*	*Chi-square*[2]
A	62.7	70.2	67.3	416	41.9
B	70.3	67.8	68.7	416	55.6

[1] A firm was predicted to be an adopting firm if $P>0.5$.

[2] The critical value of $\chi^2(1)$ at the 0.001 confidence level is 10.83.

contemporaneous return on a portfolio of firms of equivalent size. This measure is implemented by using 20 equally weighted "size portfolios" constructed by ranking all NYSE and AMEX firms with returns available on the CRSP daily returns file according to market value of common equity on December 31, 1973.[14]

---

[14] We use excess returns rather than raw returns to reduce the variance of the error term in the regression, and to reduce cross-sectional dependence in the error term.

The "unexpected earnings" variable, *UE*, is measured by subtracting the "expectation" of earnings per share at the beginning of the disclosure period from the "realization" of earnings made public during the disclosure period, adjusted to a consistent accounting basis. The "expectation" is the *Value Line* earnings forecast published on the day preceding the disclosure interval. The "realization" is either an earnings announcement or an earnings forecast that is released during the disclosure interval.[15] For all firms, *UE* is deflated by the most recent month-end stock price preceding the disclosure interval.

Although *UE* provides a complete measure of unexpected earnings for nonadopting firms, it reflects only one component of unexpected LIFO earnings for adopting firms. For these firms, unexpected LIFO earnings can be viewed as the difference between unexpected earnings on a pre-change accounting basis (captured by *UE*) and an unexpected LIFO effect. The final independent variable in equation (7), *ULE*, is included in the regression to represent the second of these components. *ULE* is measured analogously to *UE* as the difference between the "realized" and "expected" LIFO earnings effect. The "realized" LIFO effect is either an actual effect or a forecast that is released during the disclosure interval, and the "expected" LIFO effect is *HG74* (see table 2) adjusted to an after-tax basis. Like *UE*, *ULE* is deflated by the most recent month-end stock price preceding the disclosure interval.

Results

Table 6 reports estimation results for equations (6) and (7) using estimates of *P* from both model A and model B.[16] Each set of results includes coefficient estimates and *p*-values based on White's (1980) consistent covariance estimator. The *p*-values for $\hat{\gamma}_1$ reflect two-sided tests, as do those for $\hat{\gamma}_0$. However, *p*-values for $\hat{\gamma}_2$ and $\hat{\gamma}_3$ (the "earnings response coefficients") reflect one-sided tests.

The operational test of our null hypothesis is $\gamma_1 = 0$. For adopting firms, the table 6 results weakly support rejection of the null hypothesis for model B, but not for model A. In contrast, for nonadopting firms, the results very strongly support rejection of the null hypothesis for both models.[17]

The results also indicate that $\hat{\gamma}_1$, which estimates the average value of *B* in equation

---

[15] In most instances, the "realization" for adopting firms is disclosed on a LIFO basis, and the "expectation" is on a pre-change accounting basis. Therefore, for adopting firms, "realized" earnings is adjusted by the LIFO earnings effect in order to place it on a consistent accounting basis with the earnings "expectation." For nonadopting firms, the earnings "realization" and the earnings "expectation" are always on a consistent accounting basis.

[16] The estimation of equation (6) is based on 254 of the 258 nonadopting firms for which *P* is estimated. Nonadopting firms were excluded from the returns analysis if *UE* exceeded half of the firm's market value of equity (three firms) or if the firm was not traded on the day preceding its disclosure interval (one firm). The estimation of equation (7) is based on 142 of the 158 adopting firms for which *P* is estimated. Adopting firms were excluded from the returns analysis if *ULE* could not be estimated (12 firms) or if either *UE* or *ULE* exceeded half of the firm's market value of equity (four firms).

[17] There are several possible explanations for the relatively weaker results for adopting firms. First, the variance of the coefficient on ΔP may be larger for adopting firms than for nonadopting firms, so that it is more difficult to detect an average price response to disclosure for adopting firms. Second, it is possible that *P* is estimated with more error for adopting firms than for nonadopting firms, so that the estimated coefficient on $\Delta \hat{P}$ is more biased toward zero for adopting firms. Third, table 4 implies that the distribution of $\Delta \hat{P}$ is more concentrated near zero for adopting firms than for nonadopting firms. If these estimates are accurate, there is less variation in $\Delta \hat{P}$ for adopting firms than for nonadopting firms, so that it may be more difficult to reliably detect an association with excess returns for adopting firms. Finally, *B* may be nonzero for many adopting firms, but may have a mean of zero across the sample.

Table 6
Disclosure Interval Regression Results[1]

Panel A. Adopting Firms:

Model:[2] $Ret = \gamma_0 + \gamma_1 \Delta\hat{P} + \gamma_2 UE + \gamma_3 ULE + \epsilon$

	γ_0	γ_1	γ_2	γ_3	*Adjusted* R^2
Model A					
Coefficient	−0.030	0.024	0.799	−0.174	0.063
White's probability	(0.178)	(0.388)	(0.104)	(0.150)	
Model B					
Coefficient	−0.038	0.049	0.804	−0.192	0.073
White's probability	(0.063)	(0.092)	(0.101)	(0.121)	

Panel B. Nonadopting Firms:

Model:[2] $Ret = \gamma_0 + \gamma_1 \Delta\hat{P} + \gamma_2 UE + \epsilon$

	γ_0	γ_1	γ_2	*Adjusted* R^2
Model A				
Coefficient	0.032	0.070	0.110	0.032
White's probability	(0.005)	(0.003)	(0.268)	
Model B				
Coefficient	0.034	0.067	0.124	0.027
White's probability	(0.005)	(0.006)	(0.241)	

[1] For each model, the first row presents ordinary least squares coefficient estimates. The second row presents probability values based on White's (1980) covariance estimator. Probability values for γ_0 and γ_1 are based on two-sided tests, and those for γ_2 and γ_3 are based on one-sided tests. The regressions were estimated with 142 (254) adopting (nonadopting) firms.

[2] Variables in the regression are defined as follows:

Ret	Returns measured over an interval ending on the earliest identifiable disclosure date for the LIFO decision and beginning on the day following the preceding *Value Line* report date, adjusted by a portfolio return for firms of equivalent size.
$\Delta\hat{P}$	Estimated revision in probability of adoption at the disclosure of the LIFO decision, equal to $1-\hat{P}$ for adopting firms and $-\hat{P}$ for nonadopting firms.
UE	Difference between the "actual" earnings per share (adjusted for the LIFO effect for adopting firms) and the most recent *Value Line* forecast, deflated by market value of equity at the end of the month preceding the disclosure interval.
ULE	Difference between the actual earnings effect of adopting LIFO and HG74 (on an after-tax basis), deflated by market value of equity at the end of the month preceding the disclosure interval (for adopting firms only).

(4), is positive for both adopting and nonadopting firms.[18] Thus, the average price response to disclosure was positive for adopting firms (because ΔP was positive) and negative for nonadopting firms (because ΔP was negative).

---

[18] Recall from equation (3) that, for a given firm, B represents the difference between the firm's "LIFO-conditional" and "FIFO-conditional" values, as assessed by investors prior to disclosure. See appendix A, particularly equations (A3) through (A5), for a more detailed discussion of LIFO- and FIFO-conditional equity values under a set of simple assumptions about the behavior of investors and managers.

Finally, the intercept estimates reported in table 6 suggest that, during the disclosure intervals in our sample, there was a difference between mean excess returns for adopting and nonadopting firms that was not related either to disclosure of the LIFO decision or to the earnings realization. The return difference, which is -0.062 (-0.072) for model A (model B), is significant at conventional levels.[19]

Alternative Explanations

Our rejection of $\gamma_1 = 0$ is consistent with the hypothesis that stock prices changed in response to disclosures of 1974 LIFO decisions. In this section, we examine two alternative explanations for this result. First, the significant result observed for $\hat{\gamma}_1$ may be due to correlation between $\Delta \hat{P}$ and measurement error in *UE*.[20] Second, our results may be unrelated to disclosure of the LIFO decision if stock returns are systematically associated with the financial and operating characteristics that are summarized by $\Delta \hat{P}$.

To address the first possibility, we reestimated equations (6) and (7), substituting ranks of *UE* and *ULE* for the values of the unexpected earnings variables. In both regressions, the γ_1 estimates and their significance levels were generally similar. Thus, the γ_1 results reported in table 6 do not appear to be due to measurement error in *UE*.

We used two procedures to address the second possibility. First, we reestimated equations (6) and (7) for successive periods in event time during the year following the date of the LIFO disclosure.[21] Beginning 11 days after the disclosure date, we estimated each equation for each of nine consecutive, nonoverlapping 28-day return intervals, the length of the median disclosure interval. The γ_1 coefficient estimates for these nine nondisclosure intervals indicate no association between $\Delta \hat{P}$ and returns beyond that which would be expected by chance (see table 7).

Second, we held the set of sample disclosure intervals constant for each regression, and assigned the *j*th disclosure interval to a randomly selected firm *i* whose own disclosure interval did not overlap interval *j*. An excess return was then computed for firm *i* over disclosure interval *j*. This process was repeated for each sample disclosure interval and produced a sample of returns for randomly selected firms over "nondisclosure" intervals that have precisely the same distribution in calendar time as the disclosure intervals on which the table 6 results are based. These returns were regressed on the original independent variables for firms in the random sample to obtain γ_1 estimates that are not related to the disclosure of LIFO decisions. By repeating this entire procedure 1,000 times, we generated empirical distributions for γ_1 that capture the relation (if any) between $\Delta \hat{P}$ and returns during our sample disclosure intervals that is *not* due to disclosure of LIFO decisions.

These empirical distributions are used to determine whether the coefficients reported in table 6 are unusual relative to coefficients based on returns during the sample disclosure intervals for randomly selected nondisclosing firms. The results of

---

[19] Similar findings are reported by Ricks (1982). Biddle and Ricks (1988) examine whether the return difference reported by Ricks is due to differential biases in analysts' forecasts of earnings between adopting and nonadopting firms. Our analysis indicates that this difference in mean returns persists after considering the effect of disclosure of the LIFO decision and unexpected earnings.

[20] A separate question is whether our results are influenced by collinearity between *UE* and $\Delta \hat{P}$. The sample correlations are very small: -0.051 (-0.055) for adopting (nonadopting) firms with model A, and 0.002 (-0.055) for adopting (nonadopting) firms with model B. None of these is significant at any conventional level.

[21] We are grateful to a referee for suggesting this procedure.

Table 7
Estimates of the Coefficient on $\Delta\hat{P}$ for Post-Disclosure Return Intervals

Interval (in days)	Adopting Firms		Nonadopting Firms	
	Model A	Model B	Model A	Model B
(11,38)	0.013	−0.023	0.006	−0.033
	(0.546)	(0.287)	(0.809)	(0.265)
(39,66)	0.010	0.015	0.005	0.010
	(0.637)	(0.521)	(0.830)	(0.703)
(67,94)	0.044	0.003	0.008	−0.016
	(0.077)	(0.893)	(0.657)	(0.366)
(95,122)	0.014	0.007	−0.020	−0.012
	(0.587)	(0.823)	(0.266)	(0.527)
(123,150)	−0.028	−0.016	0.006	0.013
	(0.147)	(0.488)	(0.726)	(0.478)
(151,178)	−0.009	−0.023	−0.005	−0.036
	(0.711)	(0.388)	(0.784)	(0.091)
(179,206)	−0.002	−0.011	0.001	0.019
	(0.942)	(0.685)	(0.943)	(0.300)
(207,234)	−0.008	−0.019	0.001	−0.026
	(0.675)	(0.424)	(0.965)	(0.209)
(235,262)	−0.035	−0.049	0.027	−0.011
	(0.116)	(0.040)	(0.178)	(0.558)

Note: The table presents γ_1 estimates for adopting firms from the following regression model:

$$Ret=\gamma_0+\gamma_1\Delta\hat{P}+\gamma_2 UE+\gamma_3 ULE+\epsilon,$$

and for nonadopting firms from the following regression model:

$$Ret=\gamma_0+\gamma_1\Delta\hat{P}+\gamma_2 UE+\epsilon.$$

The first row presents ordinary least squares coefficient estimates. The second row presents probability values based on White's (1980) covariance estimator. Probability values are based on two-sided tests. The regressions were estimated by using 142 (254) adopting (nonadopting) firms.

Variables in the regression are defined as follows:

Ret	Returns measured over the indicated interval, adjusted by a portfolio return for firms of equivalent size. Day zero is the earliest identifiable disclosure date for the LIFO decision.
$\Delta\hat{P}$	Estimated revision in probability of adoption at the disclosure of the LIFO decision, equal to $1-\hat{P}$ for adopting firms and $-\hat{P}$ for nonadopting firms.
UE	Difference between the "actual" earnings per share (adjusted for the LIFO effect for adopting firms) and the most recent *Value Line* forecast, deflated by market value of equity at the end of the month preceding the disclosure interval.
ULE	Difference between the actual earnings effect of adopting LIFO and *HG74* (on an after-tax basis), deflated by market value of equity at the end of the month preceding the disclosure interval (for adopting firms only).

this conservative procedure are consistent with those reported earlier. For model A, the empirical p-value for adopting (nonadopting) firms is 0.468 (0.009). For model B, the p-value for adopting (nonadopting) firms is 0.146 (0.015). This indicates that the results reported in table 6 were not due to an association between $\Delta\hat{P}$ and returns that is specific to our sample disclosure intervals but unrelated to LIFO disclosures.

VI. Concluding Remarks

In an inflationary environment, a change to LIFO accounting for inventories typically reduces reported earnings, while at the same time increasing after-tax cash flows. Whether investors respond negatively to the earnings effect of the change or positively to the cash flow effect has been a question of continuing interest. In attempting to answer this question, accounting researchers have repeatedly examined the behavior of stock prices for firms adopting LIFO. However, no study to date has produced a satisfactory answer regarding the earnings effect versus the cash flow effect. Moreover, no previous study has offered convincing evidence that stock prices respond *in any way* to disclosures that LIFO has been adopted.

Our research provides new evidence on the second question. Taken as a whole, our cross-sectional regression results suggest that stock prices did respond to disclosures of LIFO decisions during fiscal 1974. However, as our analysis in appendix A demonstrates, the price response to disclosure depends on investors' prior beliefs regarding: (1) variables that affect the decision, and (2) information conveyed by the decision about other aspects of the firm's activities. Therefore, unless investors' prior beliefs can be specified, no inference can be made from our results about the cause of the price response, or about why the price response varies across firms. These questions remain to be addressed in future research.

Appendix A
Direction of Stock Price Response to Disclosure of a LIFO Decision

Section II expresses the price change at the time a firm's LIFO decision is disclosed as:

$$\Delta V_i = A_i + B_i \Delta P_i, \qquad \text{(A1)}$$

where:

$$B_i = E_0(V_i \mid \text{Adopt}, Z_0) - E_0(V_i \mid \text{Not Adopt}, Z_0), \qquad \text{(A2)}$$

and Z_0 is the set of public information prior to disclosure. The price response to disclosure of the LIFO decision itself is the second term on the right-hand side of equation (A1), $B_i \Delta P_i$. This model relies only on the assumption that investors price publicly available information.

As Kang (1988) suggests, by example, the sign of the price change in response to disclosure of the LIFO decision can be either positive or negative for both adopting and nonadopting firms, depending on investors' prior beliefs. With respect to equation (A1), this means that B_i could be either positive or negative. This can be demonstrated after making the following additional assumptions:

1. Managers know with certainty the present value of incremental income taxes (T) due to inventory holding gains if the firm does not adopt LIFO, and implementation costs (I) if LIFO is adopted.
2. Managers act in the interest of stockholders, adopting LIFO if and only if T exceeds I, and investors believe this to be the case. Thus, a disclosure that LIFO has been adopted informs investors that T exceeds I.
3. Investors are uncertain about T, I, and V^*, where V^* is the value of the firm exclusive of T and I, the direct cash flow effects of the LIFO decision. Prior to disclosure, they assess the joint distribution of these variables conditional on the set of publicly available data, denoted as Z_0.

These assumptions represent an environment in which: (1) managers have more precise information than investors, (2) managers act in the interests of stockholders, and (3) investors use publicly available information to anticipate the actions of managers.

Under these assumptions, a disclosure that LIFO has been adopted informs investors that T exceeds I, and that the firm has chosen to incur LIFO implementation costs to avoid incremental taxes on inventory holding gains. Thus, for a given firm, the expected "LIFO-conditional" equity value can be expressed as:

$$E_0(V \mid \text{Adopt}, Z_0) = E_0(V^* - I \mid T > I, Z_0). \qquad \text{(A3)}$$

Similarly, a disclosure that LIFO has *not* been adopted informs investors that I equals or exceeds T. In this case, the firm has chosen to bear incremental income tax costs (if any) because LIFO implementation costs are greater. Thus, the expected "FIFO-conditional" equity value is:

$$E_0(V|\text{Not Adopt},Z_0)=E_0(V^*-T|T\leq I,Z_0). \tag{A4}$$

Substituting equations (A3) and (A4) into equation (3), we obtain:

$$\begin{aligned} B&=E_0(V^*-I|T>I,Z_0)-E_0(V^*-T|T\leq I,Z_0)\\ &=E_0(T|T\leq I,Z_0)-E_0(I|T>I,Z_0)\\ &\quad+[E_0(V^*|T>I,Z_0)-E_0(V^*|T\leq I,Z_0)]. \end{aligned} \tag{A5}$$

Equation (A5) indicates that, for a given firm, B includes three components: (1) $E_0(T|T\leq I,Z_0)$, which is investors' prior assessment of the present value of tax savings that will be foregone if the firm *does not adopt* LIFO; (2) $E_0(I|T>I,Z_0)$, which is investors' prior assessment of the present value of implementation costs that will be incurred if the firm *does adopt* LIFO; and (3) $[E_0(V^*|T>I,Z_0)-E_0(V^*|T\leq I,Z_0)]$, which represents investors' prior assessment of additional information about firm value that will be communicated by the LIFO decision.[22]

From equation (A5), the sign of B, and therefore the sign of the price response to disclosure, cannot be predicted *a priori* without knowledge of investors' beliefs prior to the LIFO disclosure. Kang (1988) uses two examples to demonstrate that the direction of the price response to disclosure is ambiguous. First, if investors know T with certainty, but are uncertain about I, the difference between the first two components of B is positive. Second, if investors know I with certainty, but are uncertain about T, the difference between the first two terms of B is negative.

The final component of B complicates matters further. Previous studies, including Kang, have assumed that the LIFO decision conveys no information about V^*. However, this need not be the case. The disclosure of a LIFO adoption (nonadoption) decision may convey to investors the information that managers expect input costs for the firm to increase at a faster (slower) rate than had been expected. In this case, the disclosure could result in a revision in V^* that is independent of the tax consequences of the LIFO decision and that may vary from firm to firm, depending on investors' beliefs regarding the extent to which firms are able to "pass on" shifts in input costs to their customers.

In summary, even though the model described above is based on very simple assumptions about the behavior of managers and investors, the price response to disclosure of a LIFO decision is not likely to be predictable as to sign (much less magnitude), given the current state of knowledge regarding investors' prior beliefs.

Appendix B

Variables Used to Estimate the Probability of Adopting LIFO

The first two variables shown in table 2 (group 1) serve as proxies for LIFO tax savings foregone in fiscal 1973 (*HG73*) and for expected LIFO tax savings available in fiscal 1974 (*HG74*).[23] The computational algorithm for these variables is similar to that developed by Biddle (1980).[24] *INFLATION*, an annualized industry-specific measure of the rate of inventory cost change during the three months preceding the firm's disclosure interval, estimates expected future price changes.

Inventory price changes used in calculating these variables are based either on a Bureau of Labor Statistics producer price index (for manufacturers) or a Consumer Price Index (for retailers) series closely

[22] Note that, after structure is added to the model, it is apparent that the price change at disclosure of the LIFO decision is not proportional to the unconditional expectation of LIFO tax savings as proposed in previous studies (Biddle and Lindahl 1982; Stevenson 1987).

[23] In small samples examined by Morse and Richardson (1983) and Dopuch and Pincus (1988), realized inventory holding gains tend to increase in the year immediately preceding a LIFO adoption, and to exhibit a dramatic increase in the adoption year itself.

[24] *HG73* and *HG74* are calculated assuming that LIFO was adopted at the beginning of the relevant year. Under these circumstances, the holding gain is the difference between the FIFO and LIFO values for the ending inventory. In the case of *HG73*, a dollar value LIFO procedure is used to estimate the LIFO cost of the ending inventory, and the FIFO cost is the firm's reported ending inventory. In the case of *HG74*, both the LIFO and the FIFO ending inventory amounts must be estimated because actual ending inventory is not known at the beginning of the disclosure interval. We assume no change in real inventory during 1974. Thus, LIFO ending inventory equals beginning inventory, and expected FIFO ending inventory is beginning inventory restated to end-of-year prices, using price series that were publicly available at the beginning of the firm's disclosure interval.

associated with the firm's primary four-digit SIC code. A four-digit SIC code price index was available for only 76 of the 309 four-digit industries (comprising 198 firms) in the sample. To extend this subsample, we matched four-digit industry descriptions with other Producer and Consumer Price Index series. Thus, we were able to estimate these variables for an additional 292 firms in 101 four-digit SIC codes. In all cases, calculations are based only on price series that were publicly available prior to the beginning of each firm's disclosure interval.

Variables included in group 2 of table 2 are financial-statement-based measures of various firm characteristics that may be associated with the probability of LIFO adoption. Some of these measures are intended to distinguish among firms according to their abilities to take advantage of potential LIFO tax benefits. For example, given an upward drift in inventory costs, a significant potential for LIFO tax benefits is more likely to exist when inventory is material in relation to overall operations (*INV, CGSSAL*), when the rate of inventory turnover is low (*CGSINV*), when inventory levels are stable (*CVINV, CLDBT*), and when the nondomestic component of inventories is small (*FOREIGN*). Likewise, firms that were consistently profitable (*ROA, ROS, CVINCOME*) and those without significant operating loss carryforwards (*NOL*) are more likely to be able to take advantage of potential LIFO tax benefits. Larger firms (*ASSET*) may be better able to take advantage of potential LIFO tax benefits because they can more easily afford the administrative burden of LIFO accounting. Moreover, large firms may be able to avoid inventory liquidations, either because LIFO can be implemented with larger, more diversified inventory pools or because specialized inventory management talents tend to be available to these firms. Similarly, it has been argued (Lee and Hsieh 1985) that relatively capital-intensive firms (*GFA*) need to devote greater resources to inventory planning and management, and may therefore adapt more readily to LIFO. Finally, since LIFO generally results in lower reported earnings, lower retained earnings, lower net tangible assets, and a diminution of current assets relative to current liabilities, a LIFO adoption may have the effect of tightening constraints imposed on a firm's activities by accounting-based covenants in existing loan agreements. Accordingly, firms with relatively more long-term debt (*LTD*) or small current assets relative to current liabilities (*CR*) are less likely to adopt LIFO.[25] Similarly, managers of firms with widely dispersed ownership may have compensation arrangements tied more closely to reported earnings than do managers of firms with concentrated "insider" ownership. If this is the case, firms with substantial inside ownership (*OWNER*) will be more likely to adopt LIFO than those with diffuse ownership.[26]

Group 3 variables are based on a reading of the *Value Line Investment Survey* throughout the study period. *Value Line* analysts appear to have devoted considerable attention, both on a firm-by-firm basis and at the industry level, to the potential effects of LIFO adoption. *VL*1 and *VL*2, represent *Value Line*'s LIFO analysis for each sample firm as of the beginning of its disclosure interval. The dummy variable *VL*1 takes a value of 1 when *Value Line* reports that the firm's management was "considering" a LIFO adoption.[27] A second dummy variable, *VL*2, takes a value of 1 when *Value Line* did not report a "considering" statement by management but did provide a discussion of the potential tax and/or earnings impacts of LIFO adoption for the firm under review. Such an analysis could indicate some likelihood of adoption. The third variable in group 3, *CHANGE*, represents *Value Line*'s expectation of earnings performance in fiscal 1974 relative to that in fiscal 1973.

Finally, group 4 contains two measures of the frequency of LIFO adoption in the firm's industry. The first variable, *IND*1, measures the pre-fiscal 1974 proportion of firms within each two-digit SIC code that used LIFO as their primary inventory method. *IND*2 measures frequency of fiscal 1974 LIFO adoptions within the firm's two-digit SIC code, on the basis of data most recently available at the beginning of each firm's disclosure interval. Both of these variables are measured at the two-digit SIC level because our sample is not large enough to meaningfully measure them for finer industry partitions.[28]

---

[25] The relation between debt-related contracting costs and LIFO adoption has been examined by Lee and Hsieh (1985), Hunt (1985), and Dopuch and Pincus (1988).

[26] The empirical results on this issue are mixed. Hunt (1985) reports a negative relation between LIFO usage and management ownership. Niehaus (1989) reports the same result for low levels of management ownership, but reports a positive relationship for high levels of management ownership. For additional discussion of the potential impact of concentrated ownership, see Smith (1976), Abdel-khalik (1985), and Healy et al. (1987).

[27] The language used in *Value Line* varied, but generally attributed to management words like "considering," "mulling," and "contemplating."

[28] Both variables were coded as 0 for firms in two-digit industries represented by fewer than 10 firms. This procedure affected 26 (19) firms for *IND*1 (*IND*2).

References

Abdel-khalik, A. R. 1985. The effect of LIFO-switching and firm ownership on executives' pay. *Journal of Accounting Research* 23 (Autumn): 427–47.

———, and J. C. McKeown. 1978. Understanding accounting changes in an efficient market: Evidence of a differential reaction. *The Accounting Review* 53 (October): 851–68.

Biddle, G. C. 1980. Accounting methods and management decisions: The case of inventory costing and inventory policy. *Journal of Accounting Research* 18 (Supplement): 235–80.

———, and F. W. Lindahl. 1982. Stock price reactions to LIFO adoptions: The association between excess returns and LIFO tax savings. *Journal of Accounting Research* 20 (Autumn): 551–88.

———, and W. E. Ricks. 1988. Analyst forecast errors and stock price behavior near the earnings announcement dates of LIFO adopters. *Journal of Accounting Research* 26 (Autumn): 169–94.

Brown, R. M. 1980. Short-range market reaction to changes to LIFO accounting using preliminary earnings announcement dates. *Journal of Accounting Research* 18 (Spring): 38–63.

Cushing, B. E., and M. J. LeClere. 1992. Evidence on the determinants of inventory accounting policy choice. *The Accounting Review* 67 (April): 355–66.

Dopuch, N., and M. Pincus. 1988. Evidence on the choice of inventory accounting methods: LIFO versus FIFO. *Journal of Accounting Research* 26 (Spring): 28–59.

Healy, P. M., S. Kang, and K. G. Palepu. 1987. The effect of accounting procedure on CEO's cash salary and bonus compensation. *Journal of Accounting & Economics* 9 (April): 7–34.

Hunt, H., III. 1985. Potential determinants of corporate inventory accounting decisions. *Journal of Accounting Research* 23 (Autumn): 448–67.

Kang, S. H. 1988. The stock price effects of a LIFO adoption. Working paper, Massachusetts Institute of Technology, Cambridge.

Lanen, W., and R. Thompson. 1988. Stock price reactions as surrogates for the net cash flow effects of corporate policy decisions in cross-sectional studies. *Journal of Accounting & Economics* 10 (December): 311–34.

Lee, C. J., and D. Hsieh. 1985. Choice of inventory accounting methods: Comparative analyses of alternative hypotheses. *Journal of Accounting Research* 23 (Autumn): 468–85.

Lindahl, F. W. 1989. Dynamic analysis of inventory accounting choice. *Journal of Accounting Research* 27 (Autumn): 201–26.

———, C. Emby, and R. H. Ashton. 1988. Empirical research on LIFO: A review and analysis. *Journal of Accounting Literature* 7: 310–33.

Niehaus, G. R. 1989. Ownership structure and inventory method choice. *The Accounting Review* 64 (April): 269–84.

Morse, D., and G. Richardson. 1983. The LIFO/FIFO decision. *Journal of Accounting Research* 21 (Spring): 106–27.

Ricks, W. E. 1982. The market's response to the 1974 LIFO adoptions. *Journal of Accounting Research* 20 (Autumn): 367–87.

———. 1986. Firm size effects and the association between excess returns and LIFO tax savings. *Journal of Accounting Research* 24 (Spring): 206–16.

Siegel, S., and N. J. Castellan, Jr. 1988. *Nonparametric Statistics for the Behavioral Sciences.* New York: McGraw-Hill.

Smith, E. D. 1976. The effect of the separation of ownership from control on accounting policy decisions. *The Accounting Review* 51 (October): 707–23.

Stevenson, F. L. 1987. New evidence on LIFO adoptions: The effects of more precise event dates. *Journal of Accounting Research* 25 (Autumn): 306–16.

Sunder, S. 1973. Relationship between accounting changes and stock prices: Problems of measurement and some empirical evidence. *Journal of Accounting Research* 11 (Supplement): 1–45.

White, H. 1980. A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity. *Econometrica* 40 (May): 817–38.

THE ACCOUNTING REVIEW
Vol. 67, No. 2
April 1992
pp. 355–386

Evidence on the Determinants of Inventory Accounting Policy Choice

Barry E. Cushing
University of Utah
Marc J. LeClere
University of Maryland

SYNOPSIS AND INTRODUCTION: This study provides additional evidence on factors influencing inventory accounting policy choice. Following Lee and Hsieh (1985), Dopuch and Pincus (1988), and Lindahl (1989), we compare long-time FIFO users with long-time LIFO users to test variables that might be expected to influence inventory method choice. This study incorporates a tax savings variable in evaluating the LIFO/FIFO choice. We also survey financial executives to corroborate the results of secondary data analyses and to seek new insights about inventory method choice.

The findings suggest that: (1) anticipated tax savings is the primary reason firms use LIFO and (2) other firms do not use LIFO because of numerous factors without a single dominant reason. Most of these factors diminish the potential tax savings from LIFO. They include LIFO layer liquidations, LIFO bookkeeping costs, declining production costs, and contradictory tax and financial reporting rules on accounting for inventory obsolescence. However, other factors include effects on debt covenants, concern about the complexity of LIFO, and the requirements of FIFO for government contracts.

Two unexpected findings emerged. First, the multivariate model is quite accurate in predicting FIFO firms, but predicts less then half of those using LIFO. Second, the correspondence between the responses to the FIFO survey and the cross-sectional data is not as strong as might be

We wish to acknowledge the helpful comments received on earlier drafts of this article from Gary Biddle, Charles Enis, Tom Frecka, Herb Hunt, Bruce Lubich, Tom Omer, Morton Pincus, Lynn Pringle, Terry Shevlin, and Joni Young, and from workshop participants at the University of Illinois, Lehigh University, and The Pennsylvania State University. We also wish to thank Alex Chen, Claudio Frumento, and Felicia LeClere for their assistance with our statistical analyses, and Martha Wartick for her assistance in preparing and distributing our survey instrument.

Submitted October 1990.
Accepted August 1991.

expected, which suggests that determinants of inventory choice continue to be elusive.

Key Words: *Inventory policy, LIFO, FIFO, Accounting choice.*

Data Availability: *Data is available from the authors upon request.*

THE remainder of this article consists of the following. Section I describes the sample selection procedures and the variables included in the cross-sectional tests. Section II presents the results of the univariate tests, the multivariate model in which LIFO/FIFO choice is the dependent variable, the survey of the chief financial officers, and a firm-by-firm comparison of survey responses and firm financial and operating characteristics that attempts to infer each FIFO firm's reasons for not using LIFO. Section III presents the summary and conclusions.

I. Sample Design and Specification of Variables

Sample of Firms

The sample is comprised of COMPUSTAT firms that used either FIFO or LIFO exclusively for all inventories during the period 1975–84.[1] To increase the likelihood of identifying the characteristics of LIFO and FIFO users, firms that changed inventory accounting methods during the test period are excluded from the sample.[2] The final sample consists of 175 FIFO firms and 48 LIFO firms.

Based on two-digit SIC codes, there are 39 industries represented among the 175 FIFO firms, and 18 industries among the 48 LIFO firms. The largest clusters of FIFO firms are in electrical machinery (31 firms), instruments (17 firms), and chemicals (12 firms). Among the LIFO firms, six are in the fabricated metal products industry, five are in the machinery industry, and no other industry contains more than four LIFO firms. Thus, no single industry dominates the sample.

Specification of Variables

From prior studies and the inventory advice offered in trade journals, several variables were identified as reported in table 1. The estimated tax savings associated with LIFO instead of FIFO is directly proportional to the difference between a firm's cost of goods sold under LIFO and FIFO, assuming the same long-term marginal tax rate for all firms.[3] The cost of goods sold differences for all sample firms for the years 1976–84 were estimated by using a dollar value LIFO algorithm. Consistent with the

---

[1] The number of LIFO firms in the sample is increased substantially by using a sample period subsequent to 1974–75, when large numbers of firms switched to LIFO. Restricting the sample to firms that use LIFO or FIFO exclusively simplifies the estimation of LIFO tax savings.

[2] Non-U.S. firms were excluded from the sample, as LIFO is generally not allowed under foreign tax rules. The sample included only firms reporting net sales, total inventory, and operating income before depreciation for each of the 10 years.

[3] Morse and Richardson (1983, 109) point out that marginal tax rates may differ because of differences in state corporate income tax rates and tax loss carrybacks or carryforwards. In defending their assumption that marginal tax rates are constant across all firms, they point out that most state corporate income tax rates are within 3 percent of each other, with a range of less than 10 percent from highest to lowest. Our study does not control for differences in state corporate income tax rates, but our design incorporates a tax loss carryforward variable.

Table 1
Variables and Operational Measures

Variable	*Authors*	*Operational Measure*
V_1 Estimated tax savings	Bd*, DP*, LP*, MR*	Ratio of LIFO/FIFO cost of goods sold difference to net sales for each year, averaged over the nine-year sample period.
V_2 Inventory materiality	DP, LH, Ld	Nine-year average of the ratio of FIFO or "as-if-FIFO" inventory to total assets.
V_3 Average tax loss carryforward	CS*, Ld*	Average of the ratio of net additions to tax loss carryforwards to operating income for each year.
V_4 Inventory variability	Bd, DP*, LH*, Nh*	Coefficient of variation of FIFO or "as-if-FIFO" inventory over the nine-year sample period.
V_5 Inventory obsolescence		Ratio of research and development expense to net sales for each year, averaged over the nine-year period.
V_6 Firm size	DP*, HZ, LH, Ld*, MR*, Nh, ZH*	Nine-year average of reported net sales.
V_7 Leverage	CS, DP, Hn*, LH, Ld*, Nh ZH*	Largest value (over the nine-year study period) of the ratio of long-term debt less capitalized lease obligations to net tangible assets.
V_8 Current ratio	Hn	Nine-year average of current assets to current liabilities.

* Indicates that the variable was found to be a significant predictor of LIFO/FIFO choice.

Index to Author Codes:

Bd = Biddle (1980)
CS = Caster and Simon (1986)
DP = Dopuch and Pincus (1988)
HZ = Hagerman and Zmijewski (1979)
Hn = Hunt (1985)
LH = Lee and Hsieh (1985)
LP = Lee and Petruzzi (1989)
Ld = Lindahl (1989)
MR = Morse and Richardson (1983)
Nh = Niehaus (1989)
ZH = Zmijewski and Hagerman (1981)

approach used by Biddle (1980) and Dopuch and Pincus (1988), our algorithm estimates cost of goods sold differences according to the addition and liquidation of inventory layers.[4]

---

[4] The first step in the algorithm is to convert the reported LIFO inventory values of the LIFO firms to FIFO by adding the excess of the current acquisition cost over the LIFO inventory value (the so-called "LIFO reserve") to the reported LIFO inventory value. Values of the LIFO reserve were obtained from firm financial statements and 10-K disclosures. For the period from 1975 to 1984, this provided a series of ending FIFO inventory values for each of the LIFO firms. These were converted to constant dollar values. The constant dollar inventory layers were then identified for each year and, where appropriate, liquidated in subsequent years. For each year, the algorithm computes the difference between the current cost of the inventory (the FIFO value) and the estimated LIFO cost, which is the sum of the cost of the inventory layers established in prior years and the current year. For any specific year, the cost of goods sold difference is the amount by which this inventory cost difference exceeds the inventory cost difference of the prior year. This procedure yields cost of goods sold differences for the nine-year period from 1976 to 1984. These cost of goods sold differences were

The algorithm was validated by comparing its estimates for the 48 LIFO firms with the changes in the LIFO reserves reported in the firms' annual report and 10-K disclosures. The ratio of the *actual* cost of goods sold difference to net sales averaged 0.0121, while the ratio of the *estimated* cost of goods sold difference to net sales averaged 0.0116. This difference in means was not statistically significant ($t = -0.256$; $p = 0.80$). In addition, the rank order correlation between actual and estimated cost of goods sold differences for the LIFO firms was 0.775 ($p < 0.01$).

Firms with low levels of inventory would expect small tax savings from using LIFO and are expected to use methods other than LIFO. Firms with high and unpredictable variability in physical inventory levels may consider LIFO undesirable because of the adverse tax and cash flow effects from unanticipated LIFO layer liquidations (Granof and Short 1984; Jannis et al. 1980). Lee and Hsieh (1985), Dopuch and Pincus (1988), and Niehaus (1989) report significantly lower inventory variability among the LIFO firms in their samples. To provide a comparable measure of inventory variability, we converted the reported inventory amounts for LIFO firms to a FIFO basis.

Firms that experience a high rate of inventory obsolescence may find LIFO undesirable because inventory obsolescence arising from the replacement of old product lines by newer products raises a difficult LIFO accounting question for which authoritative accounting and tax guidelines are lacking, and a divergence of views and interests exists between the SEC and IRS (Jannis et al. 1980, 351–55). There is evidence that the number of new inventions produced by firms is highly correlated with the size of firm research and development expenditures, after adjusting for firm size (Mansfield et al. 1977, 9). Therefore, as a proxy for a firm's propensity to experience inventory obsolescence, we used the nine-year average of research and development expense as a percentage of net sales.

Bookkeeping cost was cited as a deterrent for adopting and using LIFO (Granof and Short 1984; Hilke 1986). Estimates of LIFO adoption costs provided by the respondents to Hilke's survey of wholesalers ranged from 0.7 to 7.9 percent of average profits.[5] This wide range has led us to use firm size as a proxy for LIFO bookkeeping costs.

We include a leverage measure (long-term debt less capitalized lease obligations, divided by net tangible assets) in this study to account for the adverse effect of these methods on accounting ratios that are used for debt covenants (Watts and Zimmerman 1986).[6] The firm's largest value of this ratio over the nine-year study period was used as the operational measure of leverage.

---

scaled by net sales in each year. and averaged over the sample period. For consistency, we calculated all of our other variables over this same nine-year period. While the inflation rates used by Biddle in his algorithm were derived from wholesale price indices, we follow Dopuch and Pincus in employing producer price indices. We attempted to match COMPUSTAT's industry description for each sample firm to the producer price index for an individual commodity item. If a match could not be obtained for an industry description and commodity item, successive attempts sought a match between industry description and product class, subgroup, and group indices, respectively. Producer price indices were not employed for two sets of firms—those involved in transportation and services and those involved in wholesale and retail sales. For transportation and services firms, the consumer price index for all urban consumers was used, while for wholesale and retail firms a price index for finished goods excluding foods was employed.

[5] These are averages across several industry categories, and represent total one-time compliance costs, including conversion costs plus the present value of anticipated maintenance cost. For details, see Hilke (1986, 25).

[6] This is the reciprocal of a measure often incorporated into bond covenants (Smith and Warner 1979, 136).

Table 2
Univariate Analysis of LIFO/FIFO Choice

Variables	*LIFO Firms*	*FIFO Firms*	*t-value*	*Chi-square*
V_1 Estimated tax savings	0.0116	0.0086	−3.20**	
V_2 Inventory materiality	0.2280	0.2620	1.74	
V_3 Average tax loss carryforward	0.0127	0.0297		6.14*
V_4 Inventory variability	0.2709	0.3760	4.86***	
V_5 Inventory obsolescence	0.0070	0.0171	3.89***	
V_6 Firm size	1,514M	373M		8.81**
V_7 Leverage	0.2085	0.2852	3.97***	
V_8 Current ratio	3.2279	2.5648	−2.93**	

* Significant at 0.05.
** Significant at 0.01.
*** Significant at 0.001.
M Millions of dollars.

Also with respect to debt covenants, Wilson and Christensen (1985, 5) indicate that the lower (LIFO) balance sheet value reported for inventories could affect compliance with debt covenants based on the current ratio. To provide a measure that is comparable across firms, we computed the current ratio after adjusting inventory values for the LIFO firms to an "as-if-FIFO" basis.[7,8]

II. Results and Discussion

Univariate Tests

Table 2 presents the means and test statistics for each variable. Since the distributions for two of the variables showed evidence of extreme nonnormality, the sample differences for these variables were evaluated with chi-square tests. For the remaining variables, differences between sample means were evaluated with a *t*-test

---

[7] Another factor often cited as influencing management's selection of accounting procedures is the existence of incentive compensation plans that link management compensation to reported earnings. Empirical studies either fail to find an association between this variable and LIFO/FIFO choice (Abdel-khalik 1985; Hagerman and Zmijewski 1979; Hunt 1985) or suggest that LIFO adoptions have a small effect on CEO compensation (Healy et al. 1987). In view of this evidence, we did not include a compensation measure in our cross-sectional tests.

[8] Prior studies examined industry membership (Lee and Hsieh 1985; Lindahl 1989; Niehaus 1989) but industry membership might proxy for underlying firm characteristics that tend to vary by industry (Lindahl 1989). For this reason, we did not include industry dummy variables in our model. As a practical matter, our LIFO sample may have been too small to enable meaningful tests of industry variables. We tested several other variables that have been included, with mixed results, in previous LIFO/FIFO studies. These included frequency of price increases (Lee and Hsieh 1985; Dopuch and Pincus 1988), average inflation rate (Dopuch and Pincus 1988; Lindahl 1989), price variability (Dopuch and Pincus 1988; Lee and Hsieh 1985), natural business year (Caster and Simon 1986), and management ownership (Hunt 1985; Niehaus 1989). With minor exceptions, these variables were not significant in either our univariate or multivariate tests, so we do not report on them further. Details of these tests are available on request.

Table 3
Multivariate Analysis of LIFO/FIFO Choice

Variables	*Model 1*	*Model 2*	*Model 3*
Intercept	−5.62	−0.09	−4.23
V_1 Estimated tax savings		208.87***	168.28***
V_2 Inventory materiality		−6.86***	−7.49***
V_3 Average tax loss carryforward		−3.43	0.39
V_4 Inventory variability		−3.48*	−2.23
V_5 Inventory obsolescence		−46.72**	−54.35**
V_6 Natural log of firm size	0.59***		0.59***
V_7 Leverage	−2.88*		−3.02
V_8 Current ratio	0.63***		0.66**
R^2	.143	.170	.254
Model chi-square	39.25***	49.40***	74.95***
Degrees of freedom	3	5	8
% correctly predicted	79.8	82.1	84.8
Log likelihood function	−96.52	−91.45	−78.67

* Significant at 0.05.
** Significant at 0.01.
*** Significant at 0.001.

based upon independent samples with the assumption of unequal variances (Snedecor and Cochran 1980).

The differences are statistically significant at the 0.05 level for seven of eight variables. The results indicate that, in comparison to LIFO firms, FIFO firms have lower estimated tax savings as a percentage of sales, have larger average tax loss carryforwards, have greater variability in inventory levels, are smaller, are more highly leveraged, and have lower current ratios. All of these results are consistent with prior studies. Our inventory obsolescence proxy variable is also significant—FIFO firms have a higher ratio of research and development expense to sales, as expected. In conformance with Dopuch and Pincus (1988), inventory materiality was higher for our FIFO firms, but the difference was not significant.

Multivariate Tests

Three logistic regressions were estimated (table 3) to assess the relative effect of the variables on the firm's choice of inventory method. Model 1 included only the variables having expected effects unrelated to tax savings: firm size, leverage, and current ratio. The chi-square value for this model is significant at $p < 0.001$; all three variables are significant in the predicted direction and the R^2 level of 0.143 is relatively low.[9] This model correctly predicts LIFO/FIFO choice for 178 of the 223 firms, or 79.8 percent. But since there are 175 FIFO firms in the sample (or 78.5 percent of the total), this model's prediction accuracy is little better than a naive model predicting that all firms use FIFO.

---

[9] R^2 in a logistic regression model is called "pseudo-R^2," but its interpretation is analogous to the coefficient of determination in linear regression, and approaches 1 as the predictive power of the model improves.

Model 2 incorporates all five of the tax-related variables. The chi-square value for this model is also significant at $p<0.001$ with an R^2 of 0.170. This model correctly predicts LIFO/FIFO choice for 82.1 percent of the sample firms. Compared to model 1, these results suggest that tax considerations alone explain only slightly more of the LIFO/FIFO choice behavior of these firms than do variables reflecting nontax economic consequences of LIFO/FIFO choice. In model 2, four of the tax-related variables are significant—tax savings, inventory materiality, inventory variability, and inventory obsolescence. However, the sign of the inventory materiality variable is the opposite of that expected.

Model 3 incorporates all of the variables in models 1 and 2. The chi-square value for this model is significant at $p < 0.001$ with an R^2 of 0.254, and prediction accuracy is 84.8 percent. Six of eight variables in this model are significant at about the 0.05 level. As in model 2, the sign of the inventory materiality variable is the opposite of that expected. Average tax loss carryforward and inventory variability are not significant. Likelihood ratio tests were employed to test the significance of the incremental explanatory value of leverage and each of the significant variables in model 3. The value of the likelihood ratio was significant at the 0.05 level for all six variables, indicating that all of these variables add significant incremental explanatory value to the model.

Our results are consistent with earlier studies and do improve upon prior models, with the exception of the higher R^2 value reported by Lee and Hsieh. Perhaps most surprising is that inventory variability is not consistently significant in our multivariate models, which is in contrast to the findings of Lee and Hsieh (1985) and Niehaus (1989). Also surprising is that inventory materiality is highly significant but in the opposite direction from our expectations.

In an overall sense, the prediction accuracy of the models with respect to LIFO firms is low. Model 3 is highly accurate at predicting which firms use FIFO, predicting 166 out of 175 (94.9 percent), but rather inaccurate at predicting which firms use LIFO, predicting only 23 of 48 (47.9 percent). Examining prediction probabilities shows that, while most of the FIFO firms fit model 3's profile of a FIFO firm, many LIFO firms do not fit its profile of a LIFO firm.

The Survey

A questionnaire was sent to the chief financial officer of each of the 223 sample firms.[10] FIFO and LIFO firms each received a separate version of the questionnaire. Each questionnaire listed a number of reasons why a firm might elect to use LIFO (FIFO) rather than FIFO (LIFO) and asked respondents to indicate whether each reason was: (1) the most important reason, (2) an important reason, (3) of minor importance, or (4) unimportant or irrelevant in their selection of LIFO (FIFO). The reasons listed on the questionnaire included those related to tax considerations and other economic consequences, as discussed earlier, plus what we labeled "orthodox" reasons—reasons consistent with traditional accounting theory, such as a belief that FIFO better reflects the actual physical flow of inventories, or a desire to report on a basis consistent with other firms in the same industry.

---

[10] In the two years following the close of our sample selection period, 15 of our sample firms had been acquired. The questionnaire was sent to the last known chief financial officer of each of these firms, and responses were received from four of them (27 percent). Also, two of the firms had switched from 100 percent FIFO to 100 percent LIFO, but neither of these responded. Thirteen had switched methods for a portion of their inventories, and seven of these responded.

Usable responses were received from 32 of the 48 LIFO firms, a response rate of 67 percent, and from 70 of the 175 FIFO firms, or 40 percent. Responding and nonresponding firms were compared with respect to all of the variables used in the previous tests to test for nonresponse bias. The only significant difference found was that FIFO respondents were larger than FIFO nonrespondents ($\chi^2=4.94$; $p=0.026$). On that basis, we concluded that there is little evidence of significant nonresponse bias in our survey results.

The responses to the LIFO questionnaire, summarized in panel A of table 4, indicate that LIFO's capacity to lower a firm's income taxes is the most important reason for its use. The orthodox reason that LIFO provides a better matching of costs and revenues was also considered important by most respondents. Industry conformity was cited as important by a few respondents, while concerns about political visibility associated with higher levels of reported earnings under FIFO were of little concern to the respondents.

Responses to the FIFO questionnaire are summarized in panel B of table 4. The two reasons ranked most important were the orthodox accounting reasons that FIFO better reflects the actual physical flow of inventories and provides a more accurate inventory valuation on the balance sheet. Interestingly, all of the reasons listed on the questionnaire were cited as important by some of the respondents. In addition, several other reasons were identified by respondents in written comments, including declining unit production costs due to cost reduction and productivity improvement programs (five respondents), a feeling that LIFO is too complex for staff or management to understand (three respondents), requirements that FIFO be used for costing on certain government contracts (two respondents), fast turnover of new products (two respondents), and the presence of substantial income from foreign subsidiaries for which LIFO is not allowed (one respondent).

Comparison of Cross-sectional and Survey Results

Responses to the LIFO survey appear to refute the "size hypothesis." This suggests that the association between firm size and LIFO observed in secondary data may be spurious. Additional analyses were performed for FIFO respondents using the "most important" or "important" factors. The results are summarized in table 5. Three significant differences are observed. First, the firms of the 11 FIFO respondents citing "immaterial inventory" had a significantly lower inventory-to-assets ratio than other FIFO firms. Second, the firms of the 21 FIFO respondents citing concern about LIFO bookkeeping costs had significantly lower average net sales than other FIFO firms. Both of these results are consistent with expectations. However, the firms of the 13 FIFO respondents citing concern about LIFO layer liquidations had significantly lower inventory variability than other FIFO firms, which is contrary to expectations. It is possible that inventory variability is a poor proxy for concern about LIFO layer liquidations.

Survey responses and variables examined in the cross-sectional tests were also

Table 4
Analysis of Survey Responses

Panel A. Reasons for Preferring LIFO to FIFO:

Possible Reasons for Using LIFO	*Response Code Number*[a] 1	2	3	4	*Avg.*[b]
Tax/Cash Flow Reasons					
Lowers income taxes	27	5	0	0	1.16
Other Economic Consequences					
Increased reported earnings under FIFO could lead to unfair charges of excess profits	0	0	3	29	3.91
Orthodox Reasons					
Provides a better matching of costs and revenues	6	19	5	2	2.09
Conforms with other firms in the same industry	0	4	10	18	3.44

Panel B. Reasons for Preferring FIFO to LIFO:

Possible Reasons for Using FIFO	*Response Code Number*[a] 1	2	3	4	*Avg.*[b]
Tax/Cash Flow Reasons					
LIFO would not provide tax benefits	20	15	6	29	2.63
Concern about LIFO bookkeeping costs	4	17	18	31	3.09
Concern about possible liquidation of LIFO layers	6	7	14	43	3.34
Immateriality of inventories	9	2	5	54	3.49
Other Economic Consequences					
Concern about the effects on stock prices of reporting lower earnings under LIFO	13	14	9	34	2.91
Concern that debt covenants are based on FIFO-based accounting numbers	1	4	7	58	3.74
Orthodox Reasons					
Better reflects the actual physical flow of inventories	20	19	13	18	2.41
Provides a more accurate inventory valuation in the balance sheet	11	32	11	16	2.46
Conforms with other firms in the same industry	5	12	11	42	3.29

[a] Per the questionnaire, response codes indicate the following:

1 = The most important reason why we use LIFO (FIFO).
2 = Not the most important reason why we use LIFO (FIFO), but an important reason, or at least an important advantage of using LIFO (FIFO).
3 = Though an advantage of LIFO (FIFO), this reason is of minor importance to our choice of LIFO (FIFO).
4 = Unimportant or irrelevant to our choice of LIFO (FIFO).

Note that three respondents did not follow directions and marked some reasons with an "X" while leaving others blank. In these cases, "X" was coded as 1, and "blank" as 4. Eleven other respondents coded some responses as instructed while leaving others blank; in these cases blanks were also coded as 4.

[b] Avg. is a weighted average of the responses.

Table 5

Comparison of Survey Responses and Related Variables

		Average Value of Variable for FIFO Respondents Citing As		
Survey Item	*Related Variable*	*Important* (n_1)	*Other* (n_2)	*Test Statistic*
No tax benefits ($n_1=35$; $n_2=35$)	Estimated tax savings	0.0084	0.0081	$t=0.23$
	Tax loss carryforward	0.0561	0.0042	$\chi^2=1.87$
	Inventory obsolescence	0.0160	0.0191	$t=-0.41$
Immaterial inventory ($n_1=11$; $n_2=59$)	Inventory/assets ratio	0.0396	0.3131	$t=-13.26^{**}$
LIFO layer liquidations ($n_1=13$; $n_2=57$)	Inventory variability	0.2698	0.3755	$t=-2.36^{*}$
Bookkeeping cost ($n_1=21$; $n_2=49$)	Firm size	292M	635M	$\chi^2=8.23^{**}$
Debt covenants ($n_1=5$; $n_2=65$)	Debt/equity ratio	0.2977	0.2961	$t=0.03$
	Current ratio	2.7105	2.5111	$t=0.58$

* Significant at 0.05.
** Significant at 0.01.
M Millions of dollars.

compared firm by firm for each of the 70 FIFO respondents. This analysis identified one or more matches of reasons for 33 of the 70 FIFO firms responding to the survey.[11] On the basis of extended comparisons, likely reasons for using FIFO rather than LIFO were identified for 15 additional firms, as reported in table 6.

A total of 12 different factors were identified as likely reasons why firms used FIFO rather than LIFO, though all but four of these reasons are related to a lack of significant tax benefits from LIFO.[12] Of the 22 firms for which no likely reasons for using FIFO could be identified, three switched to LIFO for part or all of their inventories during the three years after fiscal 1984, which was the close of our sample selection period. However, this still leaves 19 firms, or 27 percent of the respondents, for which we could not identify any convincing explanation for the continuing use of FIFO rather than LIFO.

III. Summary and Conclusions

Reasons for inventory choice continue to be elusive. Several factors have been identified and were examined here with secondary data and survey results. Prediction

---

[11] Firms were considered to be "matched" if important reasons for FIFO use cited on the questionnaire corresponded to important reasons for FIFO use indicated by the cross-sectional data. For those variables for which low values would be expected to be associated with the use of FIFO (such as estimated tax savings), the 10th percentile value of that variable for all LIFO firms in the sample was used as a threshold. When a firm fell short of this threshold on that variable, the variable was identified as an important reason for that firm's use of FIFO. An analogous procedure was used for variables for which high values would be expected to be associated with the use of FIFO.

[12] The four exceptions are: concern about LIFO bookkeeping costs, concern about debt covenants, concern about LIFO complexity, and requirements for using FIFO for government contracts.

Table 6
Most Likely Reasons for Use of FIFO Rather Than LIFO Among Survey Respondents

Likely Reason	*Among 33 Matched Firms*	*Among Other Firms*	*Total by Reason*
Set 1: *Anticipated Reasons*			
Immaterial inventory	10		10
Low estimated tax savings	9	5	14
Inventory obsolescence	9		9
Tax loss carryforwards	7	2	9
Concern over LIFO bookkeeping cost	3		3
Concern over LIFO liquidations	2		2
Concern over debt covenants	2		2
Set 2: *Other Reasons*			
Declining production costs	2	3	5
Publishing industry	1	2	3
LIFO too complex	1	2	3
Government contracts	1	1	2
Significant foreign operations		1	1

of LIFO continues to be difficult, although survey respondents point to tax savings as the primary factor. Biddle (1980) raised the issue of why many firms seemingly forego millions of dollars in potential tax savings by not switching from FIFO to LIFO. The answer consists of many different reasons, most of which are relevant to only a small fraction of FIFO firms. Uncertainty about the future behavior of these various factors would also tend to inhibit firms from switching to LIFO, given that such a switch is not easily reversed (Lee and Petruzzi 1989). Our results suggest that both tax and nontax factors play an important role.

References

Abdel-khalik, A. R. 1985. The effect of LIFO-switching and firm ownership on executives' pay. *Journal of Accounting Research* 23 (Autumn): 427–47.

Biddle, G. C. 1980. Accounting methods and management decisions: The case of inventory costing and inventory policy. *Journal of Accounting Research* 18 (Supplement): 235–80.

Caster, P., and D. T. Simon. 1986. The association between selected variables and inventory methods. *Quarterly Review of Economics and Business* 26 (Summer): 84–93.

Dopuch, N., and M. Pincus. 1988. Evidence on the choice of inventory accounting methods: LIFO vs. FIFO. *Journal of Accounting Research* 26 (Spring): 28–59.

Granof, M. H., and D. G. Short. 1984. Why do companies reject LIFO? *Journal of Accounting, Auditing, & Finance* 7 (Summer): 323–33.

Hagerman, R. L., and M. E. Zmijewski. 1979. Some economic determinants of accounting policy choice. *Journal of Accounting & Economics* 1 (August): 141–61.

Healy, P. M., S. Kang, and K. G. Palepu. 1987. The effect of accounting procedure changes on CEOs' cash salary and bonus compensation. *Journal of Accounting & Economics* 9 (April): 7–34.

Hilke, J. C. 1986. Regulatory compliance costs and LIFO: No wonder small companies haven't switched. *Journal of Accounting, Auditing, & Finance* 1 (Winter): 17–29.

Hunt, H. G., III. 1985. Potential determinants of corporate inventory decisions. *Journal of Accounting Research* 23 (Autumn): 448–67.

Jannis, C. P., C. H. Poedtke, Jr., and D. R. Ziegler. 1980. *Managing and Accounting for Inventory.* New York: John Wiley & Sons.

Lee, C. J., and D. A. Hsieh. 1985. Choice of inventory accounting methods: Comparative analysis of alternative hypotheses. *Journal of Accounting Research* 23 (Autumn): 468–85.

——, and C. R. Petruzzi. 1989. Inventory accounting switch and uncertainty. *Journal of Accounting Research* 27 (Autumn): 277–96.

Lindahl, F. W. 1989. Dynamic analysis of inventory accounting choice. *Journal of Accounting Research* 27 (Autumn): 201–26.

Mansfield, E., J. Rapoport, A. Romeo, E. Villani, S. Wager, and F. Husic. 1977. *The Production and Application of New Industrial Technology.* New York: W. W. Norton & Company.

Morse, D., and G. Richardson. 1983. The LIFO/FIFO decision. *Journal of Accounting Research* 21 (Spring): 106–27.

Niehaus, G. R. 1989. Ownership structure and inventory method choice. *The Accounting Review* 64 (April): 269–84.

Smith, C. W., Jr., and J. B. Warner. 1979. On financial contracting: An analysis of bond covenants. *Journal of Financial Economics* 7 (June): 117–61.

Snedecor, G. W., and W. G. Cochran. 1980. *Statistical Methods.* Ames: The Iowa State University Press.

Watts, R., and J. Zimmerman. 1986. *Positive Accounting Theory.* Englewood Cliffs, NJ: Prentice-Hall.

Wilson, P. W., and K. E. Christensen. 1985. *LIFO for Retailers: A Business, Financial, and Tax Guide.* New York: John Wiley & Sons.

Zmijewski, M. E., and R. L. Hagerman. 1981. An income strategy approach to the positive theory of accounting standard setting/choice. *Journal of Accounting & Economics* 3 (August): 129–49.

SHORTER ARTICLES

1. Charles D. Bailey and Edward V. McIntyre
   Some Evidence on the Nature of Relearning Curves

2. Stephen K. Asare
   The Auditor's Going-Concern Decision: Interaction of Task Variables and the Sequential Processing of Evidence

3. Dan Givoly and Carla Hayn
   The Valuation of the Deferred Tax Liability: Evidence from the Stock Market

4. Richard P. Brief and Raef A. Lawson
   The Role of the Accounting Rate of Return in Financial Statement Analysis

5. Eugene A. Imhoff, Jr., and Gerald J. Lobo
   The Effect of Ex Ante *Earnings Uncertainty on Earnings Response Coefficients*

THE ACCOUNTING REVIEW
Vol. 67, No. 2
April 1992
pp. 368–378

Some Evidence on the Nature of Relearning Curves

Charles D. Bailey
University of Central Florida
Edward V. McIntyre
The Florida State University

SYNOPSIS AND INTRODUCTION: A number of formal models have been used to reflect the fact that people generally require less time to perform a complex task after they acquire some familiarity and experience with the task. These models are referred to as learning curves, and a sizable literature exists on their properties and uses (Yelle 1979). In contrast, little has been written on the nature of relearning curves, which are formal models used to estimate the reduction in task time that occurs while *relearning* skills that have been forgotton due to interruption. The relatively few studies addressing this aspect of interrupted production advocate "backing up" the original learning curve (typically a log-linear model) in some fashion and using it to model the relearning process (Adler and Nanda 1974; Carlson and Rowe 1976; Cherrington et al. 1987; Cochran 1968; Hoffmann 1968; Keachie and Fontana 1966; Lippert 1976).

This article reports on a laboratory study designed to provide some evidence on the nature of relearning curves. Subjects were paid a realistic wage to assemble structures with Erector Set parts. They repeated the task for approximately four hours, and three forms of marginal-time learning curves were fit to their performance data. After breaks ranging from seven to 175 days, they repeated the construction projects. The two sets of times were used to compute a measure of skill decrement and to fit nine forms of learning curves. Issues of interest are: (1) similarity of the functional forms of the best-fitting relearning and learning curves, (2) whether backing up the learning curve is a reasonable method of modeling relearning, and (3) effects of skill decrements on the relative goodness-of-fit of relearning curves.

Grants from The Florida State University Foundation provided payments to subjects. Julia Karcher and Michael Nibbelin contributed substantially to the successful conduct of the experiments, by their services as research assistants. Duane Meeter gave valuable statistical consultation, and the Ideal Division of CBS donated experimental materials. Participants in the Accounting Research Colloquium at FSU provided many useful comments on an earlier version of the article.

Submitted April 1990.
Accepted October 1991.

The results show that the best-fitting relearning curves are not of the log-linear form commonly used to model learning, but are of a form that is nonlinear on both a log-log and an arithmetic scale. Further, this new form developed in our study also provided a better fit than the log-linear model when applied to learning (pre-interruption) data. Additionally, backing up the best-fitting learning curve or starting anew with this same curve form also provided a good-fitting relearning curve model; however, backing up the classical log-linear model did not fit the data as well as other models. Finally, differences between the goodness-of-fit of the various models became more pronounced with the increase in amount of skill decrement.

These findings should be of interest to those who use learning curves to estimate labor time, especially under conditions of interrupted production.

Key Words: *Relearning curve, Learning curve, Forgetting, Estimating labor time.*

Data Availability: *The data are available for confirmation purposes and for academic research that the authors may approve. Please contact Charles Bailey.*

THE remainder of this article is organized as follows: section I develops and describes the models tested; section II explains the methods employed; section III presents the results; and section IV discusses the results and limitations.

I. Models Tested

Several formal functions can be employed to model the reduction in time associated with either a learning or relearning process. In this article, we test three learning curve models and nine relearning curve models. Two criteria were applied in the selection of both types of models: we modeled the incremental unit time rather than the cumulative average time because of the widespread application of the former compared to the latter (Liao 1988), and we restricted our models to those that can be transformed into a linear equation for estimation purposes. The specific models employed are listed in table 1 and described below.

Learning Curve Models

The three learning curve models tested are labeled L1, L2, and L3. Model L1 is the conventional log-linear learning curve, where y represents incremental unit time, a is the initial unit time, and x is production. By convention, the learning rate is expressed as 1 minus the percentage reduction in unit time that is expected every time production doubles. The parameter b is the log of the learning rate divided by the log of 2.

Inspection of the data suggested that for many subjects the relationship between unit time and repeated iterations of the task was not log-linear, which led to models L2 and L3. For these subjects, the same general curvilinear pattern was observed whether y and x were plotted on an arithmetic or log-log scale. Model L2 is the log-log analogue of the classical log-linear model. It is curvilinear on both an arithmetic scale and scale [log (x), log (y)], but is linear on scale [log-log (x), log-log (y)]. The parameter b still is

Table 1
Models Tested

Description	*Model*	*Model Number*
Learning Curves:	$y=ax^b$	L1
	$\log y=a[\log(x+1)]^b$	L2
	$\log y=ax^b$	L3
Relearning Curves:		
Category 1 (starting anew)	$y=ax^b$	R1
	$\log y=a[\log(x+1)]^b$	R2
	$\log y=ax^b$	R3
Category 2 (backing up to actual restart time)	$y=ax^b$	R1A
	$\log y=a[\log(x+1)]^b$	R2A
Category 3 (backing up to estimated restart time)	$y=ax^b$	R1B
	$\log y=a[\log(x+1)]^b$	R2B
Category 4 (continuing down the learning curve)	$y=ax^b$	R1C
	$\log y=a[\log(x+1)]^b$	R2C

defined as the log of the learning rate divided by the log of 2, but now the learning rate, R, represents 1 minus the percentage reduction in $\log y$ when $\log(x+1)$ doubles. Alternatively, the model implies that every time $x+1$ is squared, y is reduced to its previous value raised to the power R. Using $[\log(x+1)]^b$ rather than $(\log x)^b$ is necessary for model L2 to be defined at $x=1$. Finally, an estimate of time for the first unit is obtained by letting $x=1$ and solving for y, i.e., $y=e^{a(\log 2)^b}$.

Model L3 is suggested as an alternative to L2.[1] It also is curvilinear on both arithmetic and log-log scales, and it is linear for log-log (y) and $\log(x)$. The parameter b is interpreted the same for model L3 as for L1 and L2, but the value of R in L3 represents 1 minus the percentage reduction in $\log y$ expected when x doubles. Alternatively, model L3 implies that every time x doubles, y is reduced to its previous value raised to the power R.

Relearning Curve Models

The nine relearning curve models used fall into four categories. In category 1 the same functional forms as learning curves L1, L2, and L3 are applied to post-break data. That is, we begin anew by letting $x=1$ for the first post-break iteration, and determine parameter values by using linear regression.

---

[1] We also explored a translog function of the form $\log y=\beta_0+\beta_1(\log x)+\beta_2(\log x)^2$, but it does not appear to be useful in this setting. Both models L1 and L2 are special cases of this more general model, and, if $\beta_2=0$, the translog function reduces to model L1. In stepwise regressions (for both learning and relearning), β_2 was statistically significant ($\alpha=0.05$) for only a few subjects. Further, for most curves, β_2 was positive and produced an immediate, nonnormative upward trend in predicted marginal times.

Category 2 includes two models that back up two of the learning curves. Model R2A backs up curve L2, the best-fitting learning curve. We also included a test of backing up the classical learning curve, L1, using model R1A.

Several ways can be used to back up learning curves. The models in category 2 backed up the two learning curves by using the original learning curves to find the values of x (to the nearest integer) that corresponded to the actual restart time (the time required for the first post-break iteration). If x turned out to be 5, we labeled the first post-break iteration number 5, the second number 6, and so on. We used these values of x and the actual post-break values of y to determine a new set of parameter values for each relearning curve. Thus, we have not literally backed up the existing learning curves because we did not use the original learning curve parameter values. Our choice of the modified backing-up procedure was based on our desire to obtain the best-fitting relearning curves.

Category 3 relearning curve models use a different backing-up procedure. While models R1A and R2A used the *actual* restart times to determine the value of x assigned to the first post-break unit, models R1B and R2B used the *estimated* restart times for this purpose. The estimated restart times are those predicted by models R1 and R2 for the first post-break unit.

Category 4 relearning curves "continue down" the original learning curves L1 and L2. For example, if a subject performed ten iterations of the task before the break and six after the break, a new set of parameter values was determined by using the values of x from 11 to 16 and the actual post-break values of y. Again, we are not literally continuing down the original learning curve because we are not using the original parameter values. If forgetting is a significant factor affecting post-break times, one would not expect that models continuing down the original learning curves would fit as well as other relearning curve models.

II. Method

Bailey's original study (1989) provided data used in the present study to calculate 31 sets of learning and relearning curves. We also conducted a second experiment, replicating the original results in a somewhat different manual assembly task. Thirty subjects, including 11 who had participated in the first experiment, completed the second experimental task.

Procedures and Subjects

The tasks in both experiments used Erector Set parts and instructions, which we laid out at work stations on table tops to simulate a factory setting. The two experiments differed in the exact nature of the assembly task: task 1 was to assemble a "radar station," while task 2 was to assemble a "dump truck." At every work session, one of the authors or a research assistant assumed the role of foreman, answering workers' questions and pointing out gross errors. Subjects understood the importance of generating accurate data, and therefore avoided conversation and other distractions during work. The researcher/foreman recorded times to the nearest second when subjects started and when they announced completion.

Most subjects in both experiments were students, and all were of normal working age.[2] They were compensated at a wage rate of more than $4 an hour and appeared, as far as the researchers could tell, well-motivated throughout the experiment.

Measure of Goodness-of-fit

To estimate parameters, we have first linearized the models in their parameters and then applied ordinary least-squares regression. In linear form, the dependent variables differ across the models; thus, the meaning of R^2 for the regressions also would differ. Accordingly, we computed R^2 for each model using the estimates of marginal time obtained from each model (see Kaplan 1982; Kvålseth 1985).[3] For statistical comparisons, we first apply Fisher's r-to-z transformation because raw r scores are not distributed normally.

Definition of Proportion Forgotten

To distinguish a relearning environment from a learning environment, some forgetting must have occurred. Moreover, we expect that the best-fitting relearning curve will vary with the degree of forgetting. For example, if forgetting is minimal, a downward continuation of the existing learning curve may be the best model. Alternatively, if forgetting is near total,[4] a fresh start with a new curve may work best. The measure of forgetting that we employ (*PROLST*, for proportion lost) is:

$$PROLST=\sum_{i=1}^{4}(MT_i-ET_i)/(ET_1-ET_{m+1}), \tag{1}$$

where:

i = sequence number of a unit produced after the break,
MT_i = marginal time for iteration i,
ET_i = learning curve (model L2) estimate of marginal time for unit i based upon pre-break performance, and
m = number of iterations before the break.

The summation of differences in the numerator represents a measure of excessive labor time as a result of forgetting, and it is summed over four iterations (because all subjects completed at least four iterations after the interruption). The denominator is a measure of total learning achieved; thus *PROLST* is a ratio of lost time to total pre-break learning.

Hypotheses

The main research question is whether the level of forgetting affects the relearning curve shape. This leads to our first hypothesis:

---

[2] The number of nonstudent subjects was too small to allow meaningful comparison with student subjects, but the two groups were about the same average age and displayed no noticeable differences in performance.

[3] The computation of R^2 was based on the following equation, which Kvålseth (1985) describes as the "best" of eight possible equations for this statistic:

$$R^2=1-[\Sigma(y-\hat{y})^2/\Sigma(y-\bar{y})^2].$$

[4] It is doubtful that memory traces ever diminish to the extent that they would not facilitate relearning. See, e.g., MacLeod (1988).

Table 2
Performance Data

	Mean	*SD*	*Minimum*	*Maximum*
Task 1 (31 subjects):				
Initial assembly in minutes	48.12	15.66	26.62	83.80
Fastest time achieved before interruption	20.79	6.62	11.63	38.10
Learning curve % (model L1)	73.8	9.0	55.2	85.1
Repetitions before interruption	7.19	4.06	3	17
Repetitions after interruption	6.17	1.44	4	10
Length of break (days)	54.23	33.89	7	114
Proportion of learning lost	0.369	0.238	0.012	1.135
Task 2 (30 subjects):				
Initial assembly in minutes	28.67	7.97	17.27	48.23
Fastest time achieved before interruption	12.49	3.80	7.17	21.53
Learning curve % (model L1)	75.8	6.8	58.8	89.2
Repetitions before interruption	7.20	2.09	4	11
Repetitions after interruption	6.7	1.93	4	11
Length of break (days)	59.23	41.00	14	175
Proportion of learning lost	.548	.322	.139	1.224

$H_0$1: No interaction exists between the goodness-of-fit of the curves (*CURVE*) and the proportion of learning lost (*PROLST*).

An obvious, implicit hypothesis is that not all viable curve candidates will perform equally well:

$H_0$2: The mean goodness-of-fit of all nine curves are equal.

Another question of interest is the effect of the level of forgetting on the proportion of unexplained variance in the relearning data; i.e., whether the curves fit better, on average, after a little or a lot of forgetting:

$H_0$3: The mean goodness-of-fit for subjects experiencing either high or low forgetting is equal, when averaged across all nine curves.

III. Results

Table 2 summarizes the subjects' performance data.[5] We first compared the

---

[5] Most relearning curves exhibited a lower starting time and a slower learning rate than the related learning curves.

Table 3
Analysis of Variance

Factor	df	Mean Square	F	Tail Probability
Between-subjects:				
PROLST	1	0.17629	0.11	0.74
ERROR	59	1.63657		
Within-subjects:				
CURVE	8	0.53569	30.25	0.0000*
CURVE×PROLST	8	0.13783	7.78	0.0028*
ERROR	472	0.01771		

* After appropriate Greenhouse-Geisser adjustment, because the sphericity assumption is not met.

goodness-of-fit of the learning curve models to determine the choice of curve for purposes of computing the degree of forgetting (*PROLST*). Curve L2 ranks as best (mean $R^2=0.897$), L3 second (mean $R^2=0.880$), and L1 third (mean $R^2=0.876$). A repeated-measures ANOVA for differences between mean fits is highly significant ($F(2;120) = 9.12$, $p<0.01$). All individual differences between ranks also are significant in a multiple comparison (Tukey's honestly significant difference), which controlled the overall alpha level at 0.01.

To test the relearning-curve hypotheses, we used ANOVA with goodness-of-fit (r converted to Fisher's z) as the dependent variable. The two factors of interest are *CURVE* (a within-subjects factor having nine levels representing the nine relearning curves) and *PROLST* (proportion of learning lost, a between-subjects factor dichotomized at the median).

A second ANOVA included TASK, a two-level factor representing the two different tasks described above. Because TASK did not interact significantly with the other factors, which indicates that they are not contingent upon task, we collapsed data across tasks and present a joint analysis.

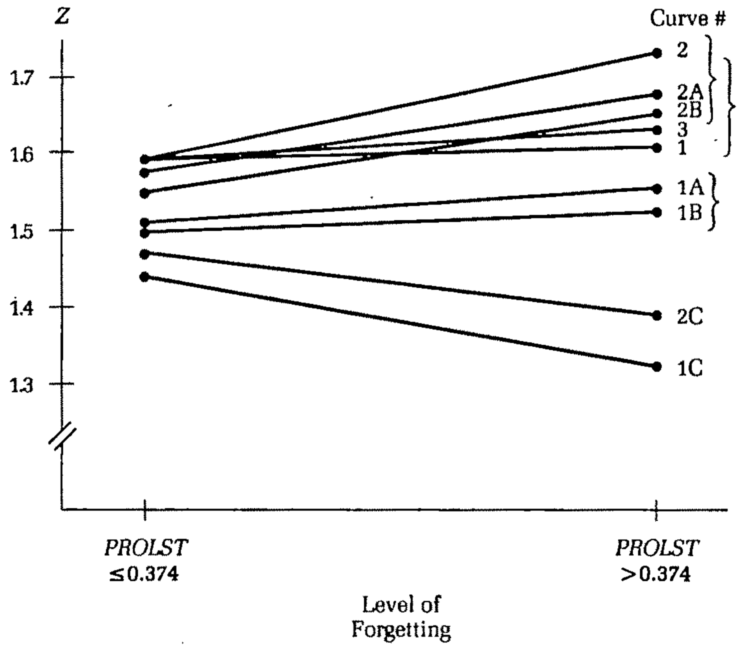
Table 3 shows the primary ANOVA results. The *CURVE×PROLST* interaction is significant ($p<0.01$), which indicates that the curve fittings are contingent upon the level of forgetting. The graphical presentation in figure 1 shows this contingency. The higher the forgetting, the more pronounced the differences between curve fits. At higher forgetting, curves 2, 2A, 2B, and 3 improve; curves 2C and 1C worsen substantially; and curves 1, 1A, and 1B are approximately unchanged. However, the relative ranking has changed only slightly.[6]

The existence of the interaction generates difficulties in the interpretations of main effects of *CURVE* and *PROLST*. Instead, it is appropriate to analyze simple effects, i.e., to investigate differences between curve fits at high, and then at low, levels of forgetting (Keppel 1982, 312). Before comparing individual curves, however, we performed one-way ANOVAs at each level to test the null hypothesis of no difference among curves. At

---

[6] Note that the average fit of the nine curves is about the same at high and low *PROLST*, which reflects the nonsignificance of the main effect of *PROLST*.

Figure 1
Interaction between Curve Fit and Level of Forgetting



Note: The brackets enclose groups of means that do not differ significantly, with α controlled for the family of tests at 0.05 by the Bonferroni method.

high *PROLST*, the test was significant ($F[8;248] = 43.34$, $p < 0.0001$), but at low *PROLST* it did not reach the 0.05 level of significance ($F[8;240] = 3.39$, $p = 0.069$). (Both probabilities are after Greenhouse-Geisser adjustments are made [Dixon 1983, 367].)

Using the Bonferroni limit (Hays 1981, 435) to control the overall significance level at 0.05, we computed *t*-tests for individual differences. Table 4 shows the results of these tests. As a visual aid to this analysis, the brackets at the right side of figure 1 connect groups of curves whose means do *not* differ significantly. For example, at the high *PROLST* level, the fits of 2, 2A, and 2B are not significantly different. Curve 2 differs significantly from curve 3, and from all other curves below curve 3 in figure 1. Curves 1A and 1B differ significantly from all other curves, but not from each other. As a final example, 1C and 2C differ significantly from all other curves.

Individual *t*-tests of difference at the low *PROLST* level also are reported in table 4. The only significant differences are between 1A and 1C, 1A and 2C, and 1B and 2C.

Supplemental Analysis

We also performed a supplemental analysis that treated *PROLST* as a continuous variable rather than as dichotomous. We plotted the difference in average fit (Fisher's z-scores) of the top five and bottom four curves $[(z_2+z_{2A}+z_{2B}+z_3+z_1)/5]-[(z_{1A}+z_{1B}+z_{2C}+z_{1C})/4]$ against *PROLST*. The R^2 of this regression is 0.17 (significant at $p < 0.01$).

Table 4
Comparisons of Curve Fits at Each Level of Forgetting

		Relearning Curve								
		1	1A	1B	1C	2	2A	2B	2C	3
RELEARNING CURVE	1		3.80	5.30	8.78	3.96	3.23	1.61	8.41	2.88
	1A	1.74		2.69	8.47	4.62	6.31	4.33	7.60	3.97
	1B	1.79	1.75		7.09	5.33	7.13	5.88	6.32	5.22
	1C	2.41	3.56	3.44		7.22	7.93	6.91	6.62	7.86
	2	0.07	1.08	1.12	1.70		1.80	2.77	6.69	4.10
	2A	0.45	2.22	2.28	2.82	0.24		2.51	7.43	2.33
	2B	1.63	1.32	1.42	2.24	0.74	2.29		6.45	0.62
	2C	2.47	4.96	4.73	1.91	1.59	3.10	2.27		7.41
	3	0.28	1.47	1.51	2.11	0.05	0.44	1.34	2.09	

Means:	1	1A	1B	1C	2	2A	2B	2C	3
r_l	.883	.873	.872	.860	.880	.883	.879	.866	.882
r_h	.901	.891	.883	.842	.914	.909	.901	.853	.903
z_l	1.569	1.506	1.503	1.441	1.571	1.562	1.543	1.469	1.572
z_h	1.616	1.549	1.521	1.314	1.721	1.680	1.652	1.367	1.640

l = low forgetting; h = high forgetting; and z = Fisher's r-to-z transformation.

Note: The numbers below the diagonal represent t-statistics for the low-forgetting group (30 df). The numbers above the diagonal represent the high-forgetting group (29 df). Underlined t-values indicate pairs of means that differ significantly after applying the Bonferroni adjustment to the family of 36 tests, i.e., the individual computed t-value must be significant at $\alpha = 0.05/36 = 0.0014$ (Hays 1981, 435).

The trend appears reasonably linear, as confirmed by a Durbin-Watson statistic of 1.84. In 45 of the 61 sets of relearning data, the top five curves fit better than the bottom four curves, as reflected by a negative difference score. This regression analysis supports our earlier conclusion that the differences in goodness-of-fit grow larger as the degree of forgetting increases.

Tests of Predictive Ability

An interesting, though limited, context exists for testing the predictive properties of our new curve versus the traditional log-log form. It is accomplished by literally backing up to a point on the original learning curve that corresponds to the initial post-break time, then using the original learning curve's parameters to project subsequent times (see, e.g., Cochran 1968; Hoffmann 1968). This test has added interest in that it represents one way in which managers can make predictions of post-break performance *before* enough observations are available to fit a relearning curve. We have used this

Table 5
Predictive Accuracy Statistics

Level of Forgetting	*MAPE*	*MSPE*
Low Forgetting (31 subjects):		
Curve L1	0.129	0.026
Curve L2	0.137	0.028
Wilcoxon signed-rank test	$z=0.80$ $p=0.20$	$z=1.10$ $p=0.46$
High Forgetting (30 subjects):		
Curve L1	0.160	0.039
Curve L2	0.133	0.028
Wilcoxon signed-rank test	$z=2.97$ $p=0.003$	$z=3.16$ $p=0.002$

Note: *MAPE* = mean absolute percentage error, and *MSPE* = mean squared percentage error.

procedure and calculated two predictive accuracy statistics, mean absolute percentage error (MAPE) and mean squared percentage error (MSPE), for both curves. The comparative results are shown in table 5.[7]

These tests are consistent with our other results. At low levels of forgetting, there is no significant difference in the average predictive ability of the two curves. At high levels of forgetting, the new curve that we suggest has significantly lower average prediction errors than the "classical" curve.

IV. Discussion and Conclusions

For purposes of discussion, the findings from our study are grouped into four categories: (1) the functional forms of the learning and relearning curve equations, (2) the ways that the relearning curve equations are applied to data, (3) the relationship between forgetting and model choice, and (4) tests of predictive ability.

1. At least for certain tasks, an equation that is curvilinear on a log-log scale (L2 and L3) provides a better fit than the traditional log-linear equation (L1). Further, since this finding applies both to learning and relearning curves, it provides some evidence that the functional form that is best for the learning curve is also best for the relearning curve.
2. Starting anew may be as effective a way to model relearning as backing up a learning curve, and backing up the classical learning curve is not necessarily the

---

[7] The error metrics are computed as follows:

$$MAPE=(1/n)\Sigma\,|(y-\hat{y})/y|,$$

$$MSPE=(1/n)\Sigma[(y-\hat{y})/y]^2,$$

where y = actual and $\hat{y}$ = predicted marginal times. On average, both curves showed a small negative bias as measured by the mean percentage error,

$$MPE=(1/n)\Sigma[(y-\hat{y})/y].$$

The average MPE across all subjects was −0.07 for curve L1 and −0.09 for curve L2.

optimal way to model relearning. Models that continued down the original learning curves (R1C and R2C) ranked last at both levels of forgetting. Further, as one would expect, the goodness-of-fit of these two models decreased as the level of forgetting increased.

3. Differences in the relative performance of the nine relearning curves become more pronounced at higher levels of forgetting, which suggests that the choice among alternative relearning curve models becomes more crucial as the level of forgetting increases. A possible explanation for these results is that at low levels of forgetting, the subjects' relearning curves are flatter and model differences are limited. At higher levels of forgetting, subjects' relearning curves initially have steeper slopes, and differences among models become more pronounced.
4. The tests of the predictive ability of literally backing up the original learning curves (i.e., using the original parameter values for L1 and L2) are consistent with our other results and provide further support for model L2 relative to the classical log-linear model.

Three limitations need to be mentioned. First, our results apply to individual learning and relearning curves, and questions remain concerning the aggregation of individual performance into group performance. Second, we were unable to test the predictive ability of the relearning curves except in a limited context. Third, our findings are based on a laboratory study that relied on student subjects, which raises questions of external validity. These issues should be of interest for future research.

References

Adler, G. L., and R. Nanda. 1974. The effects of learning on optimal lot size determination—Single product case. *AIIE Transactions* 6 (1): 14–20.

Bailey, C. D. 1989. Forgetting and the learning curve: A laboratory study. *Management Science* 35 (3): 340–52.

Carlson, J. G., and A. J. Rowe. 1976. How much does forgetting cost? *Industrial Engineering* 8 (9): 40–47.

Cherrington, J. E., S. Lippert, and D. R. Towill. 1987. The effect of prior experience on learning curve parameters. *International Journal of Production Research* 25 (3): 399–411.

Cochran, E. B. 1968. *Planning Production Costs: Using the Improvement Curve.* San Francisco: Chandler.

Dixon, W. J., chief ed. 1983. *BMDP Statistical Software.* Berkeley: University of California Press.

Hays, W. L. 1981. *Statistics*, 3d ed. New York: Holt, Rinehart & Winston.

Hoffmann, T. R. 1968. Effect of prior experience on learning curve parameters. *The Journal of Industrial Engineering* 19 (8): 412–13.

Kaplan, R. 1982. *Advanced Management Accounting.* Englewood Cliffs, NJ: Prentice-Hall.

Keachie, E. C., and R. V. Fontana. 1966. Effects of learning on optimal lot size. *Management Science* 13 (2): B102–B108.

Keppel, G. 1982. *Design and Analysis: A Researcher's Handbook.* Englewood Cliffs, NJ: Prentice-Hall.

Kvålseth, T. O. 1985. Cautionary note about R^2. *The American Statistician* 39 (November): 279–85.

Liao, S. S. 1988. The learning curve: Wright's model vs. Crawford's model. *Issues in Accounting Education* 3 (2): 302–15.

Lippert, S. 1976. Accounting for prior practice in skill acquisition studies. *International Journal of Production Research* 14 (2): 285–93.

MacLeod, C. M. 1988. Forgotten but not gone: Savings for pictures and words in long-term memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 14 (2): 195–212.

Yelle, L. E. 1979. The learning curve: Historical review and comprehensive survey. *Decision Sciences* 10 (2): 302–28.

THE ACCOUNTING REVIEW
Vol. 67, No. 2
April 1992
pp. 379–393

The Auditor's Going-Concern Decision: Interaction of Task Variables and the Sequential Processing of Evidence

Stephen K. Asare
University of Florida

SYNOPSIS AND INTRODUCTION: Sequential processing of evidence is important to auditors because it characterizes the manner in which many audit judgments are made. For example, Gibbins (1984, 110) has stated that "as a normal routine, judgment proceeds incrementally rather than by gathering full information and integrating it all before choosing a response." One advantage of sequential processing is that it allows an auditor to keep a "running total" of the effects of prior information and therefore reduces memory load (Einhorn and Hogarth 1985).

Although sequential processing of evidence promotes cognitive economy, recent findings in psychology seem to indicate that it can lead to non-normative responses, including order and framing effects (Einhorn and Hogarth 1985; Tversky and Kahneman 1981). Drawing on these general studies in psychology, previous auditing research has examined the impact of task variables (including temporal order and mode of processing) on judgments (Ashton and Ashton 1988; Butt and Campbell 1989; Tubbs et al. 1990). However, none of these studies examines how such differences in judgments are manifested in audit actions. This study extends previous research by examining whether these order effects (1) occur in alternate hypothesis frames and (2) translate into different choices of audit reports. It was hypothesized that sequential processing of evidence leads to a recency effect in auditors' going-concern judgments. This effect should lead audi-

I appreciate the criticisms and suggestions of my dissertation committee: William Waller (chairman), William Felix, and Marcia Niles. Helpful comments on earlier drafts of this article were provided by Robert Ashton, Mike Bamber, Karl Hackenbrack, Robin Hogarth, Robert Knechel, William Messier, Doug Snowball, Rick Tubbs, Arnold Wright, two unknown reviewers, and the workshop participants at the University of Georgia School of Accounting. An earlier version was presented at the 1990 Annual Meeting of the American Accounting Association. Financial support from the AICPA is gratefully acknowledged.

Submitted April 1990.
Accepted October 1991.

tors who evaluate contrary information followed by mitigating factors to issue relatively more unqualified opinions than those who evaluate the same evidence in a reverse order.

An experiment was conducted involving 70 audit managers and partners working on a case adapted from an actual audit. The auditors were randomly assigned to one of four experimental conditions created by crossing two levels of hypothesis frame and two levels of evidence presentation order. The framing manipulation involved either a failure or a viability condition. Auditors assigned to the failure (viability) frame were asked to express their belief that the firm will fail (continue in existence) during the next financial statement period. Within each framing condition, the auditors evaluated four pieces of evidence presented in two different orders. In one sequence, two mitigating factors were followed by two pieces of contrary information, and these were reversed in the other condition. After evaluating all the evidence, subjects were asked to indicate the type of opinion (audit report) they would issue.

The results support the existence of recency effects in both belief revisions and audit report choices. Auditors who evaluated contrary information followed by mitigating factors issued more unqualified (fewer modified) opinions than those who evaluated the same evidence in the reverse order. Further, as expected, hypothesis frame did not affect the existence of recency effects.

Key Words: *Order effects, Hypothesis framing, Belief revision, Reporting choices.*

Data Availability: *Data used in this study are available from the author upon request.*

THE presentation is organized as follows. The motivation for the current study is provided in the first section, which also contains an examination of key aspects of the going-concern decision, a framework for its examination, and the research hypotheses. The remaining three sections present the method, the results, and a discussion of their implications.

I. Motivation and Hypotheses Development

SAS 59 (AICPA 1990) requires auditors to evaluate the viability status of their clients as part of every audit examination and provides guidance to the auditor on steps to be followed in making such decisions. The auditor should seek and evaluate contrary information, as well as factors tending to mitigate those conditions and management plans for rectifying the underlying problems.[1] If the auditor concludes that substantial doubt exists about the client's continued existence, the audit report must reflect that conclusion.

---

[1] Contrary information may be defined as any evidence (problem) that raises doubts about the entity's ability to continue in existence, whereas mitigating factors relate to information that reduces the auditor's doubts. Thus, for subjects in the failure frame, it is expected that contrary information (mitigating factors) will lead to an increase (decrease) in beliefs. However, for subjects in the viability frame, it is expected that contrary information (mitigating factors) will lead to a decrease (increase) in beliefs.

Examining how presentation order affects going-concern decisions (audit actions) is important. Although the studies by Ashton and Ashton (1988), Butt and Campbell (1989), and Tubbs et al. (1990) examined order effects on auditors' judgments, they did not examine choices. Accordingly, their results do not provide evidence as to whether the observed differences in likelihood assessments are important in making choices. The distinction between judgment and choice is important since observed differences in judgments do not allow an inference that different actions always will result (Einhorn and Hogarth 1981; Gibbins 1984; Joyce and Biddle 1981; Lewis 1980; Schultz and Reckers 1981). Also, examining whether recency effects in auditor's belief revisions lead to different choices requires a specification of the process by which judgments (beliefs) are translated into audit choices (actions). The specification of the link between judgment and choice and the subsequent testing should contribute to our understanding of the conditions under which different judgments lead to different choices.

In evaluating a client's continuity status, an auditor may adopt either a failure or a viability hypothesis frame (Kida 1984).[2] Whether the predicted recency effects occur in both frames is an important question. Prior research suggests that a subject's hypothesis induces a confirmation bias in evidence evaluation (see, e.g., Gorman 1986). This bias causes the magnitude of belief revision to be determined by a subject's hypothesis frame, and not by the order in which evidence is received. In contrast, the belief-adjustment model (discussed in the next section) predicts a recency effect regardless of frame. By manipulating the frame, the current study seeks to extend the generalizability of the model's recency prediction and should provide some evidence on the "contrast" versus "confirmation" notions.[3]

The following aspects of the going-concern task that are instrumental to the choice of a theoretical framework and research method require further elaboration. First, as with previous research, it is assumed that the auditor forms an initial belief (anchor) about viability based on the client's financial position and updates this initial belief with the acquisition and evaluation of *additional* contrary information (bad news), mitigating factors (good news), and management plans (AICPA 1990; Kida 1980; Mutchler 1985). Although the search for and evaluation of the additional evidence occurs during the entire period of fieldwork, auditors typically obtain only a *short* series of contrary information items and mitigating factors (Mutchler 1986, 159).

Second, although in principle an auditor could wait to obtain all the evidence before evaluating (end-of-sequence evaluation), a reasonable alternative is to assume that evidence is evaluated sequentially. Sequential evaluation is more reasonable because evidence typically is received at different times from various sources and sequential evaluation is necessary to determine whether more information is required. Such conditional processing is a vital ingredient of an efficient audit (Gibbins 1984). Further, most audit partners responding to a survey regard the going-concern decision as a critical, difficult, and *complex* task (Chow et al. 1987).[4] Faced with complex tasks, decision makers were found to resort to sequential processing of evidence to ease cognitive strain (see, e.g., Hogarth and Einhorn, in press; Lopes 1987; Payne 1982; Payne et al. 1990).[5] Finally, based on interviews with practitioners, Gibbins (1984, 110) argued that, for most audit tasks, judgment "is a continuous process of receiving information, choosing to act or not act, etc."

---

[2] The audit or client factors that affect the choice of the initial hypothesis are beyond the scope of this study, which simply manipulates the hypothesis frame by giving some subjects a failure hypothesis and others a viability hypothesis.

[3] A "confirmation" behavior implies that the magnitude of a subject's belief revision is determined by the hypothesis frame, and not by the order in which evidence is presented. Conversely, a "contrast" behavior implies that subjects put more weight on inconsistent evidence that is presented later in a sequence irrespective of the hypothesis frame.

Sequential processing of evidence is also consistent with the recommendation of SAS 59 that the auditor should first evaluate conditions (including any mitigating factors) that cast doubt on the client's continuity status, and then obtain information about management plans to mitigate those conditions. Sequential processing and the related belief revisions described above refer to the auditor's judgment of viability, but the ultimate decision to modify/not modify is made after all the evidence has been evaluated. Thus, the decision process can be characterized as consisting of two stages: a sequential judgment formation followed by an end-of-sequence choice of report.

Third, the nature of belief updating in going-concern decisions involves evaluation rather than estimation (Hogarth and Einhorn, in press). The distinction between estimation and evaluation depends on the form of evidence used in belief updating. In evaluation tasks, evidence must be explicitly interpreted as confirming or disconfirming the hypothesis under consideration before a judgment is made. Further, the directional implication of evidence (confirming vs. disconfirming) is independent of the subject's current beliefs (Lopes 1987). In estimation tasks, however, the form of the evidence, which often is presented in summarized numerical form, requires little interpretation, and the directional implication of evidence depends on the subject's current beliefs. Since evidence used in the going-concern decision must be interpreted as contrary or mitigating prior to belief updating, the going-concern decision can properly be considered an evaluation task.[6]

To summarize, the going-concern decision can be characterized as a two-stage process: A judgment stage that entails an *evaluation* task in which the auditor forms an initial belief (anchor) about the client's financial distress or stability and updates this anchor *sequentially* with the acquisition and evaluation of a *short* series of *complex* evidence, and a decision stage in which the auditor decides on the type of report to issue. The belief-adjustment model proposed by Hogarth and Einhorn (in press) entails the characteristics of the judgment stage.[7] A key feature of the model is the contrast assumption implying that disconfirming evidence leads to a greater discounting of beliefs when the preceding anchor is large, but that confirming evidence leads to a greater upward revision of belief when the anchor is small. The contrast assumption thus leads to order effects in belief revisions.

---

[4] Not only is it assumed that the going-concern task is complex (cf. Chow et al. 1987) but also that it entails evaluating a complex series of evidence. Although the distinction between a complex task and a complex series of evidence is not well articulated in either the psychology or accounting literature, it is reasonable to assume that task and evidence complexity are related.

[5] The notion that cognitive strain forces decision makers to evaluate evidence sequentially rather than end-of-sequence is hardly new (see, e.g., Weld and Roff 1938, 624). Sequential processing of evidence also is consistent with the SAS 59 requirement that the effect of evidence be considered in the aggregate. Because sequential processing allows an individual to keep a running total of the effects of prior information, a subject's belief at any point in time reflects the aggregate effect of all prior evidence.

[6] Hogarth and Einhorn (in press) present evidence that estimation tasks result in data that fit averaging models, while data from evaluation tasks fit additive models. The distinction is important since the order-effect predictions depend on whether the task involves evaluation or estimation. As an example of an estimation task, consider an auditor who estimates the likelihood that a pending lawsuit against a client will result in an unfavorable outcome. The auditor might consult five legal experts, three of whom suggest that an unfavorable outcome will result. Whether this information leads to an increase or a decrease in the auditor's belief would depend on the auditor's prior belief.

[7] For a detailed specification of the model, see Hogarth and Einhorn (in press), Ashton and Ashton (1988), or Asare and Messier (1991). The model makes different order-effect predictions on the basis of various task characteristics, but this article focuses only on the order-effect prediction pertaining to recency and sequential belief revision.

Order Effects

The belief-updating model predicts a recency effect when an auditor processes a sequence of mixed evidence (contrary information and mitigating factors). Recency implies that greater weight is given to confirming evidence when it is processed after disconfirming evidence rather than before. Recency effects are presumed to be largest when both positive and negative evidence are strong, and when beliefs are highly sensitive to evidence (Hogarth and Einhorn, in press).

Support for the prediction of recency effects in auditors' belief revisions has been provided by Ashton and Ashton (1988) in an internal control context and by Tubbs et al. (1990) in an account balance estimation context. Whether these effects generalize to the current context is an appropriate question for three reasons. First, a question arises as to whether recency effects found in auditors' belief revision lead to different reporting choices; accordingly, it must be established that these recency effects in belief revision occur in the current task setting. That is, observing a recency effect in auditors' belief revisions is a necessary condition for addressing whether recency effects in judgments lead to different actions. Second, the current context involves belief revisions by more experienced auditors. Whether this strengthens or weakens the recency effects reported in previous studies is of interest. Third, an auditor could adopt either a failure or a viability frame in the current context, and it is important to examine whether the frame interacts with order effects. Prior literature suggests that the frame induces a confirmation bias in evidence evaluation (see, e.g., Fiske et al. 1983; Gorman 1986; Skov and Sherman 1986). This bias would suggest that subjects in the failure (viability) frame attach more weight to contrary information (mitigating factors) regardless of the order in which it is received. However, the contrast assumption of the belief-adjustment model implies the frame should not affect order effects, an implication that has not previously been tested. By manipulating the frame, the current study provides evidence on the generalizability of the model's recency prediction and provides some evidence on whether the recency effect or the confirmation bias is more dominant (see Ashton and Ashton 1990; Butt and Campbell 1989).

The above discussion leads to the following hypothesis.[8]

H1: The mean belief revisions of auditors who evaluate contrary information followed by mitigating factors will be larger than those of auditors who evaluate the same evidence in the reverse order.

Viability Judgments, Substantial Doubt, and Reporting Choices

The question of how judgments translate into choices has been of interest in the decision-making literature (Einhorn and Hogarth 1981). Slovic et al. (1982, 28) reviewed studies of the relationship between judgment and choice and concluded that "the information-processing strategies used prior to making choices are often quite different from the strategies employed in judgments." Specifically, they argued that single numerical judgments are prone to anchoring and adjustment processes, whereas choices are more prone to context effects that develop as a result of justification processes. Consequently, "judgments and choices often tend to be inconsistent with one another" (1982, 22). Generalizing from the findings in these studies to the going-concern setting could be problematic, however, since those studies dealt with risk-taking tasks (e.g., choices among gambles) and the going-concern decision is a hypothesis-testing task.

---

[8] Because this hypothesis assumes a viability frame, contrary information (C) is "negative" (disconfirming) evidence and mitigating information (M) is "positive" (confirming) evidence. Of course, if a failure frame is assumed, the contrast assumption implies that belief revisions of subjects who evaluate evidence in the MMCC (− − + +) sequence is larger than those of subjects who evaluate evidence in the CCMM (+ + − −) sequence. In summary, the contrast assumption implies that auditors who evaluate evidence in the CCMM sequence will assess a higher (lower) likelihood of viability (failure) than those evaluating evidence in the MMCC sequence.

Other psychology studies examining the relationship between judgment and choice have focused on jurors' decision making (see, e.g., Ostrom et al. 1978; Pennington and Hastie 1981). These studies typically characterize the juror's task as consisting of two stages. First, the juror hears conflicting testimony from a variety of sources and integrates the information into a single likelihood that the defendant is guilty. Second, the juror compares this subjective likelihood to a threshold of reasonable doubt. If the subjective likelihood of guilt is greater than the threshold, the juror's decision will be to convict the defendant, and vice versa (see, e.g., Marshall and Wise 1975; Ostrom et al. 1978; Thomas and Hogue 1976). Marshall and Wise (1975) report that 80 percent of their 158 juror subjects rendered verdicts that were consistent with this two-stage process.

In a similar fashion, the going-concern modification decision is characterized as a two-stage process. In the first stage, the auditor collects and evaluates evidence, E, in the form of ratios, contrary information and mitigating factors, to reach a subjective belief, $P(C \mid E)$, where C is the firm's continued existence. In the second stage, the auditor compares $P(C \mid E)$ to the threshold ($P^*(C)$) at which the auditor will have *substantial doubt* about the entity's ability to continue in existence (see SAS 59). If $P(C \mid E) < P^*(C)$, the auditor would have substantial doubt about the entity's ability to continue in existence and should issue a modified report; otherwise, a standard unqualified report would be issued. This discussion leads to the second hypothesis:

H2: After evaluating all pieces of evidence, auditors whose viability judgments fall below their threshold judgments will issue modified reports, and those whose viability judgments fall above or are equal to their threshold judgments will issue standard unqualified reports.

Order Effects in Reporting Choices

If the two-stage process describes the way auditors make going-concern decisions, then order effects in judgments would affect the auditor's modification decisions. Specifically, since auditors who evaluate contrary information followed by mitigating factors are hypothesized to assess a higher probability of continued existence ($P(C \mid E)$), it is expected that, controlling for $P^*(C)$, auditors in this condition will issue relatively more unqualified audit reports. This reasoning leads to the following hypothesis:

H3: Auditors who evaluate contrary information followed by mitigating factors will issue more (fewer) unqualified (modified) opinions than those who evaluate the same evidence in the reverse order.

It is important to maintain the distinction between this hypothesis and hypothesis H1 above. For example, auditors may reach different judgments about the going-concern status of a client, but still be consistent in their choice of an audit report. This could happen if they have different thresholds.

II. Method

Subjects

Representatives of four Big Six CPA firms in the Los Angeles area were provided with 120 research instruments and were asked to distribute them to managers and partners. Audit managers and partners were chosen because they are either responsible for, or provide the greatest input to, going-concern decisions. Seventy completed instruments (58%) were returned. The respondents had a mean of 12 years of auditing experience and had worked on an average of 42 audits in the preceding four years.

Design

There were two independent between-subjects variables. One concerned the order in which evidence was presented; 40 respondents were presented with two pieces of contrary information followed by two pieces of mitigating factors (CCMM), and 30 respondents had the same pieces of evidence presented to them in the reverse order (MMCC). The other independent variable was the initial hypothesis frame. Thirty-six (34) respondents were in the failure (viability) frame.

Task and Procedures

Participants were asked to assume they were reviewing a current year's audit engagement and to consider the going-concern question in conjunction with this review. The case was taken from the files of an audit engagement completed by a local office of a Big Six firm and was prepared with the close cooperation of the engagement manager and partner. The experimental materials consisted of three parts, and subjects were asked to go through the materials in the order given.

In the first part, each participant was provided with background information on a hypothetical client described as a manufacturer and marketer of teleconferencing hardware and software products for business and government. The client had incorporated 17 years earlier and its stock has been publicly traded for the most recent six years. Although the client had neither emphasized nor established a strong network of internal controls, top management monitored operations closely, and overall management controls served as adequate substitutes for detailed controls. Only minor adjustments had been made to unaudited balances in the past.

Subjects were informed that an unqualified report had been issued for the past two years that their firm had audited the client. Subjects also were told that in the past year, various members of management had expressed their differences of opinions to other employees, and that this had contributed to employee morale problems, including turnover.

After being provided with the above information, subjects were presented with the financial statements of the client for the past four years, including the current year. The client had incurred a net loss in the current and immediately preceding year, and retained earnings had a debit balance at current year-end. Financial statement balances consisted of audited values. Participants were told to consider the type of opinion to issue for the current year. After studying the background information and the financial statements, participants were asked to provide an initial judgment (on a response scale from 0–100) of the firm's likelihood of continuing in existence (or failing) through the end of the next financial statement year. This part of the experimental materials served the dual purpose of familiarizing subjects with the client's operations and financial condition, and obtaining an initial anchor (S0) for each subject.[9]

Subjects were assigned to two hypothesis-frame groups by redefinition of the response scale. For the viability frame, the lower end of the scale, "0," was labeled "certain not to continue" and the upper end, "100," was labeled "certain to continue." For the failure frame, the "0" end was labeled "certain not to fail" and the "100" end was labeled "certain to fail."

The second part of the experimental materials required subjects to make a series of judgments (S1, S2, S3, S4) based on four pieces of additional evidence available in the working papers. The specific pieces of evidence were chosen for two reasons. First, each is identified by SAS 59 as either a contrary information item or a mitigating factor. Second, the engagement partner and manager of the audit on which the materials are based indicated that these items were crucial factors in their report choice and were of approximately equal importance to the going-concern decision.[10] Each piece of evidence was presented on a separate page and participants provided a revised assessment on an appropriately labeled 0–100 scale at the end of the page. The mitigating and contrary items were presented in MMCC and CCMM orders.[11] Within the two conditions, the order of each piece of evidence was held constant. After reviewing all four pieces of evidence, participants indicated the type of opinion they would issue.[12]

Participants in the viability (failure) condition also indicated a threshold level ($P^*(C)$) below (above) which they would have substantial doubt about the entity's ability to continue in existence. To examine possible carry-over effects between $P^*(C)$ and $P(C \mid E)$, some subjects ($n=36$) were asked to provide their estimates of $P^*(C)$ after their report choices, while others ($n=34$) were asked to make that assessment prior to the S0 judgment. This manipulation was fully crossed with the frame and order manipulations.

[9] Subjects were not given an initial anchor because of the possibility of confounding source credibility and belief revision issues. Instead, subjects were allowed to form their own anchors, referred to as S0.

[10] This indication was subsequently corroborated by asking 15 partners not taking part in the experiment to rate the strength of each piece of evidence on a scale of 0–100 (very weak to very strong). The mean ratings for the contrary information items were 53.47 and 54.53, respectively (paired $t=0.399$, $p \leq 0.6953$). The mean ratings for the mitigating factors were 51.70 and 49.27, respectively (paired $t=0.707$, $p \leq 0.4911$). Finally, a one-way repeated ANOVA on the four mean ratings was not significant ($F=0.93$, $p \leq 0.4331$).

[11] Precautions were taken to ensure that the evidence was reviewed in the order provided, by having participants tear a seal to access each item of evidence. They were instructed not to tear the next seal until completely reviewing the present evidence and making a judgment. Additionally, once a particular evidence item had been reviewed and a judgment made, participants were asked not to alter previous responses. A second experiment (involving 30 managers and partners) was administered in person to provide more experimental control over the sequence in which subjects evaluated evidence. Recency effects similar to those presented in the text were also found in this second experiment. Thus, with a tighter control over the sequence, the results are unchanged. Also, post-experimental discussions with these subjects indicated that the four pieces of evidence were viewed as approximately equally important.

[12] Subjects were given three reporting options: a disclaimer, an unqualified report, or a modified report. None of the subjects chose a disclaimer. The actual audit on which the case was based had resulted in a modified report. The client did remain in existence for a year after the date of the financial statements.

Finally, all participants completed a debriefing questionnaire that asked for information about their backgrounds and auditing experience. Specifically, they were asked to indicate their primary area of responsibility in their firms, the number of audits on which they had worked in the preceding four years, and how many of those audits had resulted in unqualified, qualified, or adverse opinions. They were also asked to classify each piece of evidence presented in the experiment as being either contrary or mitigating.

A cover letter to each participant emphasized the importance of working independently and assured subjects of confidentiality. Also, the cover letter reminded participants of the SAS 59 requirement to evaluate whether substantial doubt exists about the entity's ability to continue in existence for one year beyond the date of the financial statements.

III. Results

Manipulation Check on Evidence Types

The four pieces of evidence were decomposed into seven items, and subjects classified each item as contrary or mitigating.[13] The incorrect response was chosen only 13 out of a possible 490 times.[14] Thus, the subjects' perceptions of the evidence were consistent with the intended manipulation.

Preliminary Analysis

To conduct tests that include all data, responses of subjects in the failure frame were transformed from $P(NC \mid \cdot)$ to $P(C \mid \cdot)$ by subtracting the indicated response from 100. With the transformed data, table 1 presents summary statistics on the respondents' initial assessments (S0), the paths they followed to reach their final assessments (S4), their indicated thresholds, and opinion choices. Following the response paths from S0 to S4, notice that the mean belief revisions are in the expected direction. This is further evidence on the manipulation of the sign of evidence. To control for any between-group differences in initial judgments, the dependent variable employed for statistical analysis was the difference between the initial and final judgments.

A three-way ANOVA was performed to examine whether the independent variables affected the $P^*(C)$ judgments. Timing of the $P^*(C)$ judgments was included as an independent variable to examine possible carry-over effects between $P^*(C)$ and $P(C \mid E)$. As expected, none of the independent variables affected the $P^*(C)$ judgment ($F=0.93$, $p \leq 0.4312$), indicating that neither order of evidence presentation nor hypothesis frame affected threshold judgments. This suggests that observed differences in opinion choice cannot be attributed to between-group differences in $P^*(C)$. Furthermore, the mean (standard deviation) of the $P^*(C)$ for subjects providing it *after* a report choice was

---

[13] The seven items are (1) loss of a principal customer, (2) bank does not renew line of credit and negotiations for alternatives have been unsuccessful, (3) deferment of accounts payable over a three-year period, (4) loss of a patent, (5) overhead costs are lower than for the same time last year and somewhat lower than estimates in the budget, (6) slow inventory turnover, and (7) forecast of increasing profits.

[14] One subject misclassified all responses. However, a review of this subject's belief changes revealed that all revisions were in the expected direction. Of the six remaining misclassifications, two subjects did not alter their beliefs when they reviewed the pieces of evidence that they had misclassified, three subjects altered their beliefs in the appropriate direction, and only one subject altered his or her belief in the inappropriate direction. The results of separate analyses that excluded these "violating" subjects were similar to those presented in the text.

Table 1
Descriptive Statistics on Opinion Choices, Threshold Judgments, and Viability Judgments[a]
(Means with standard deviations in parentheses)

Order	*Frame*	*N*	*N**	*S0*	*S1*	*S2*	*S3*	*S4*	*P*(C)*[b]	*Unqualified Opinion*
CCMM[c]	Fail	21	10	59.29 (21.71)	42.80 (23.92)	32.67 (21.12)	52.57 (22.75)	55.52 (20.35)	54.33 (18.97)	7
MMCC[d]	Fail	15	9	64.67 (17.97)	78.20 (14.39)	81.47 (14.73)	54.47 (19.40)	46.93 (24.52)	54.20 (18.18)	5
CCMM	Continue	19	7	70.79 (19.10)	54.84 (21.46)	46.26 (20.81)	56.26 (21.89)	69.16 (19.02)	50.53 (13.32)	13
MMCC	Continue	15	8	70.60 (19.48)	82.27 (18.97)	82.40 (11.99)	55.73 (18.45)	40.00 (13.83)	63.00 (21.03)	4

Note: N* represents subjects in each group who made their threshold assessments prior to stating their S0.

[a] Represents auditors' assessment of the likelihood (S0, S1, S2, S3, S4) that the client will continue in existence.

[b] Represents auditors' minimum threshold for the likelihood of going concern necessary to issue an unqualified report.

[c] Subjects evaluated two pieces of contrary information followed by two pieces of mitigating factors.

[d] Subjects evaluated two pieces of mitigating factors followed by two pieces of contrary information.

56.56 (16.65) and that of subjects providing $P^*(C)$ *before* S0 was 53.62 (19.63). The timing of the judgment did not affect threshold judgments ($t=0.6734$, $p \leq 0.5031$). Therefore, the results presented in the next section combine the $P^*(C)$ judgments of subjects elicited at different times. Finally, across conditions, 41 auditors chose a modified report and 29 selected an unqualified report.

Hypotheses Tests

Testing hypothesis H1 requires taking into account the framing manipulation. An analysis of variance (ANOVA) with the difference between the final (S4) and initial (S0) judgments as the dependent variable was performed. The order in which evidence was evaluated and the framing of the initial hypothesis were between-subject factors. The hypothesis frame was included to examine whether there is an interaction between frame and order. The results are presented in table 2.

A significant main effect for order is found. Neither the hypothesis frame nor the interaction of frame and order was significant at conventional levels.[15] With respect to

---

[15] To examine whether the "confirmation" or the "contrast" notion dominates when contrary information is evaluated, an ANOVA was performed with the absolute difference in likelihood rating after processing contrary information as a dependent variable. The absolute difference is S2 − S0 for subjects in the CCMM sequence and S4 − S2 for subjects in the MMCC sequence. Order of evidence presentation and the hypothesis frame were between-subject factors. The order of evidence presentation was significant ($F=8.53$, $p \leq 0.0048$), but neither the hypothesis frame ($F=0.25$, $p \leq 0.6181$) nor the interaction of frame and order ($F=1.28$, $p \leq 0.2619$) was significant. Similar analysis with mitigating evidence as a dependent variable indicated that the order of evidence presentation was significant ($F=4.74$, $p \leq 0.0330$), but both the frame ($F=0.30$, $p \leq 0.5878$) and the interaction of frame and order ($F=0.41$, $p \leq 0.5249$) were insignificant. For both types of evidence, the presentation order results were consistent with recency effects. Taken together, these results indicate that the "contrast" notion dominates the "confirmation" notion.

Table 2
Analysis of Variance: Order × Frame with the Difference between the Initial and Final Judgments (S4 − S0) as the Dependent Variable

Source	Sum of Squares	df	Mean Square	F	Probability
Order	7893.653	1	7893.653	19.70	0.0001
Frame	493.480	1	493.480	1.23	0.2712
Order × Frame	962.865	1	962.865	2.40	0.1259
Error	26448	66	400.739		

$R^2 = 0.257$.

Table 3
Subjects Categorized by the Actual Opinion Chosen and the Opinion Indicated by Their Threshold Judgments

Opinion According to Threshold	Opinion Chosen	
	Unqualified	Modified
Unqualified	28*	10
Modified	1	31*
Total	29	41

* Indicates agreement between threshold and actual opinion issued.

order, the mean belief revision (S4–S0) for subjects in the CCMM condition was −2.750 and that of subjects in the MMCC was −24.167. The main effect for evidence order and the pattern in belief change are consistent with the model's prediction of a recency effect ($F = 19.70$, $p \leq 0.001$).[16]

To test hypothesis H2, the final likelihood rating ($P(C \mid E)$), threshold judgment ($P^*(C)$), and opinion issued were compared. An auditor's modification descision was considered consistent with the two-stage reporting process if the final likelihood rating was below the threshold judgment and a modified report was chosen. Table 3 reports results of that comparison.

Subjects' modification decisions are largely consistent with the two-stage reporting process. Of the 32 subjects whose threshold judgments and final likelihood ratings indicated that they should choose a modified opinion, only one chose an unqualified opinion. However, ten of the 38 auditors whose $P^*(C)$ and $P(C \mid E)$ indicated they

---

[16] The underlying rationale for this recency effect is the contrast assumption that contrary information (mitigating factors) will induce larger (smaller) discounting (accretion) when the preceding anchor is large. This assumption was tested by correlating subjects' initial beliefs with their belief reduction (accretion) when they evaluated contrary information (mitigating factors). The results of Pearson correlations ($r = 0.42403$, $p = 0.0003$, and $r = -0.58127$, $p = 0.0001$, one-tailed tests) indicated associations consistent with this contrast assumption.

Table 4

Subjects Categorized by the Order of Evidence Presentation and Opinion Chosen

Order of Evidence Presentation	Opinion Chosen: Unqualified	Opinion Chosen: Modified
CCMM	20	13
MMCC	8	18
Total	28	31

CCMM = Auditors evaluated two pieces of contrary information followed by two mitigating factors.
MMCC = Auditors evaluated two mitigating factors followed by two pieces of contrary information.

should choose an unqualified opinion actually chose the inappropriate opinion.[17] Across all subjects a chi-square test indicated that subjects' opinion choices are consistent with the two-stage reporting process ($\chi^2 = 35.64$, $p \leq 0.001$), which is consistent with hypothesis H2.

The test of hypothesis H3 involved three steps. The first step was to check for any systematic differences in $P^*(C)$ for subjects in the two order conditions. The mean threshold judgments for auditors in the CCMM and MMCC sequence were 52.5 and 58.6, respectively, which were not significant ($t = 1.4005$, $p \leq 0.16$). Thus, the alternative explanation that any observed differences in opinion decisions are due to differences in threshold judgments is not likely.

The second step involved a consistency check to examine whether auditors' opinion choices were consistent with the relation of their final judgments (S4) to their threshold assessments. This check filters subjects whose reporting choices were not consistent with the hypothesized two-stage reporting process. Table 3 reveals that the responses of 59 subjects were consistent, but 11 were not.[18] For the 59 consistent responses only, table 4 presents descriptive statistics concerning hypothesis H3.[19]

Finally, a test of proportions indicated that auditors in the CCMM had a greater propensity to issue unqualified opinions ($Z = 2.278$, $p \leq 0.016$). Thus, the evidence is consistent with the hypothesis that recency effects will be manifest in the subjects' opinion choice.

---

[17] A proportions test for these 38 auditors indicated that the proportion of respondents choosing the appropriate (i.e., unqualified) opinion was significantly different from 0.5 ($Z = 2.92$, $p \leq 0.0018$). Possible reasons for the inconsistent responses are considered in the discussion.

[18] Further analysis indicated that these 11 subjects were not concentrated in any of the experimental cells. Specifically, five subjects assessed their $P^*(C)$ prior to providing S0 and six provided their $P^*(C)$ after they had indicated the type of opinion. With respect to the frame-by-order cells, five subjects evaluated evidence in the Fail/CCMM condition, and two subjects were in each of remaining experimental cells. Finally, all other judgments of these 11 subjects were thoroughly reviewed and appeared to be appropriate.

[19] It is hypothesized that the auditor will issue a modified report if $P(C \mid E) < P^*(C)$ and an unqualified report if $P(C \mid E) \geq P^*(C)$. Since the opinion decisions of these 11 auditors were not consistent with this notion, their responses were excluded for purposes of testing hypothesis H3. When these auditors are included in the analysis, the results are unchanged ($Z = 1.68$, $p \leq 0.0465$).

IV. Discussion

The finding that opinion choices were systematically affected by evidence order has important practical and theoretical implications. First, when evaluating management plans (e.g., cash flow projections, waivers of debt covenants, etc.), auditors may be influenced by the order in which the plans are presented to them. Several partners who participated in the second experiment recalled audits for which they had planned to modify the report but changed those plans when the client obtained a new line of credit or waiver from a key creditor at the "last minute." To mitigate this potential bias, it would seem useful for audit reviewers to evaluate evidence in a different sequence from that used by the reviewee. The finding that subjects made reporting decisions by comparing their likelihood of viability with their stated thresholds suggest that evidence order may affect the amount and type of evidence collected. Further research on this possibility is warranted.

The finding that recency occurs in both failure and viability frames also implies that *experienced* auditors do not exhibit a confirmatory strategy in hypothesis testing. This finding is not consistent with psychological studies on confirmation proneness in hypothesis testing. An interesting research issue would be to examine why auditors are prone to one form of bias (recency effects) but not to another (confirmation bias). The results also indicate that 84 percent of the auditors appeared to make modification decisions consistent with the hypothesized two-stage process. Nonetheless, some subjects' reporting decisions (16%) were not consistent with this framework. Two possible explanations for this inconsistency are that subjects may forget their thresholds at the point of choice (Fischhoff and Beyth-Marom 1983, 242) and that thresholds (or even $P(C \mid E)$) may be a range of values rather than a point estimate. These possibilities require further investigation.

Two caveats and related issues for future research need to be mentioned. First, in practice, auditors search for and evaluate evidence concurrently, and perform the going-concern task contemporaneously with other tasks. Pennington and Hastie (1986) suggest that experts faced with such dynamic tasks may construct a causal model comprising key events and relations among them. Whether the reported recency effects are mitigated by the use of such causal reasoning is an area of further research.

The second caveat is that the potential for reward and punishment was not present in this experimental setting, and it is unknown whether feelings of accountability to the firm and client existed. To the extent that the absence of rewards and accountability alter auditors' judgments and decisions (cf. Ashton 1990; Payne et al. 1990; Tetlock 1983), care must be exercised when generalizing the results of this study to audit practice.

Appendix

Mitigating and Contrary Items Presented to Subjects

Beyond 1988, the company is forecasting greatly increasing profits reaching $2 million before federal income taxes in 1992. According to recent industry statistics, the number of users of the product of the specific type distributed by the company is approximately doubling each year. By the end of February 1988, the company was already near the break-even point. Overhead costs incurred to date were substantially lower than costs for the same time last year, and somewhat lower than the estimates in the budget. **(MITIGATING ITEM)**

In February 1988, the company was successful in obtaining the deferment of $300,000 of accounts payable over a three-year period. Additionally, the company's creditors had given the company a letter indicating they would not consider the loan in default because of the late payment of the 1987 installments. The 1988 installments were not due until later in the year. **(MITIGATING ITEM)**

Subsequent to year-end, the company's bank informed management that it had no desire to renew the company's line of credit in the amount of $1,000,000. Negotiations with other banks regarding a line of credit have not been successful. **(CONTRARY ITEM)**

The company designed a new product line in mid-1987 (primarily software inventory) for approximately $400,000 (included in year-end inventory). The company has experienced problems with the quality of this product line. Inventory turnover of this product has been extremely slow, approximately once per year, and the product faces potential obsolescence because of this. Several production problems have been encountered in trying to improve the quality of this product line, and these problems have led to the loss of a principal customer. To dispose of these items, there is a need for the company to seek new customers. Also, the company's patents for the Dadex expires November 30, 1988, and this may have an adverse effect on the company's operations. As of balance sheet date, a great portion of inventory (approximately four-tenths of total assets) was either not paid for or pledged as collateral for outstanding debt. **(CONTRARY ITEM)**

References

American Institute of Certified Public Accountants. 1990. The auditor's consideration of an entity's ability to continue as a going concern. *Statement on Auditing Standards No. 59*. New York: AICPA.

Asare, S. K., and W. F. Messier. 1991. A review of audit research using the belief-adjustment model. In *Auditing: Advances in Behavioral Research*, edited by L. Ponemon and D. R. Gebhart. New York: Springer-Verlag, 111–35.

Ashton, A. H., and R. H. Ashton. 1988. Sequential belief revision in auditing. *The Accounting Review* 63 (October): 623–41.

Ashton, R. H. 1990. Pressure and performance in accounting decision settings: Paradoxical effects of incentives, feedback, and justification. *Journal of Accounting Research* 28 (Supplement): 148–80.

——, and A. H. Ashton. 1990. Evidence-responsiveness in professional judgment: Effects of positive versus negative evidence and presentation mode. *Organizational Behavior and Human Decision Processes* 46 (June): 1–19.

Butt, J. L., and T. L. Campbell. 1989. The effects of information order and hypothesis-testing strategies on auditors' judgments. *Accounting, Organizations and Society* 14 (December): 471–80.

Chow, C., A. McNamee, and D. Plumlee. 1987. Practitioners' perceptions of audit step difficulty and criticalness: Implications for audit research. *Auditing: A Journal of Practice & Theory* 6 (Spring): 123–33.

Einhorn, H. J., and R. M. Hogarth. 1981. Behavioral decision theory: Processes of judgment and choice. *Journal of Accounting Research* 19 (Spring): 1–41.

——, and ——. 1985. A contrast/surprise model for updating beliefs. Working paper, University of Chicago, IL.

Fischhoff, B., and Beyth-Marom. 1983. Hypothesis evaluation from a Bayesian perspective. *Psychological Review* 90 (July): 239–60.

Fiske, S., D. Kinder, and W. M. Larter. 1983. The novice and the expert: Knowledge-based strategies in political cognition. *Journal of Experimental Social Psychology* 19 (July): 381–400.

Gibbins, M. 1984. Propositions about the psychology of professional judgment in public accounting. *Journal of Accounting Research* 22 (Spring): 103–25.

Gorman, M. 1986. How the possibility of error affects falsification on a task that models scientific problem solving. *British Journal of Psychology* 77 (February): 85–96.

Hogarth, R. M., and H. J. Einhorn. In press. Order effects in belief updating: The belief-adjustment model. *Cognitive Psychology.*

Joyce, E., and G. Biddle. 1981. Anchoring and adjustment in probabilistic inference in auditing. *Journal of Accounting Research* 19 (Spring): 120–45.

Kida, T. 1980. An investigation into auditors' continuity and related qualification judgments. *Journal of Accounting Research* 18 (Autumn): 506–23.

———. 1984. The impact of hypothesis-testing strategies on auditors' use of judgment data. *Journal of Accounting Research* 22 (Spring): 332–40.

Lewis, B. 1980. Expert judgment in auditing: An expected-utility approach. *Journal of Accounting Research* 18 (Autumn): 594–602.

Lopes, L. L. 1987. Procedural debiasing. *Acta Psychologica* 64 (February): 167–85.

Marshall, C. R., and J. A. Wise. 1975. Juror decisions and the determination of guilt in capital punishment cases: A Bayesian perspective. In *Utility, Probability, and Human Decision Making*, edited by D. Wendt and C. Vlek. Dordrecht, Holland: Reidel Publishing Company, 257–69.

Mutchler, J. 1985. A multivariate analysis of the auditor's going-concern opinion decision. *Journal of Accounting Research* 23 (Autumn): 668–82.

———. 1986. Empirical evidence regarding the auditor's going-concern opinion decision. *Auditing: A Journal of Practice & Theory* 6 (Fall): 149–63.

Ostrom, T. M., M. J. Saks, and C. Werner. 1978. An integration theory analysis of jurors' presumptions of guilt or innocence. *Journal of Personality and Social Psychology* 36 (April): 436–50.

Payne, J. W. 1982. Contingent decision behavior. *Psychological Bulletin* 92 (September): 382–402.

———, J. R. Bettman, and E. J. Johnson. 1990. The adaptive decision-maker: Effort and accuracy in choice. In *Insights in Decision Making*, edited by R. M. Hogarth. Chicago: University of Chicago Press, 129–53.

Pennington, N., and R. Hastie. 1981. Juror decision-making models: The generalization gap. *Psychological Bulletin* 89 (March): 246–87.

———, and ———. 1986. Evidence evaluation in complex decision making. *Journal of Personality and Social Psychology* 51 (August): 242–58.

Schultz, J. J., and P. M. Reckers. 1981. The impact of group processing on selected audit disclosure decisions. *Journal of Accounting Research* 19 (Autumn): 482–501.

Skov, R., and S. Sherman. 1986. Information-gathering processes: Diagnosticity, hypothesis-confirmatory strategies, and perceived hypothesis confirmation. *Journal of Experimental Social Psychology* 22 (March): 93–121.

Slovic, P., B. Fischhoff, and S. Lichtenstein. 1982. Response mode, framing, and information-processing effects in risk assessment. In *New Directions for Methodology of Social and Behavioral Science: Question Framing and Response Consistency*, edited by R. M. Hogarth. San Francisco: Jossey-Bass, 21–36.

Tetlock, P. E. 1983. Accountability and the perseverance of first impressions. *Social Psychology Quarterly* 46 (December): 285–92.

Thomas, E. A., and A. Hogue. 1976. Apparent weight of evidence, decision criteria and confidence ratings in juror decision making. *Psychological Review* 83 (November): 442–65.

Tubbs, R., W. F. Messier, and W. R. Knechel. 1990. Recency effects in the auditor's belief revision process. *The Accounting Review* 65 (April): 452–60.

Tversky, A., and D. Kahneman. 1981. The framing of decisions and the psychology of choice. *Science* 211 (January): 453–58.

Weld, H. P., and M. Roff. 1938. A study in the formation of opinion based upon legal evidence. *American Journal of Psychology* 51 (October): 609–28.

THE ACCOUNTING REVIEW
Vol. 67, No. 2
April 1992
pp. 394–410

The Valuation of the Deferred Tax Liability: Evidence from the Stock Market

Dan Givoly
Tel-Aviv University and Northwestern University
Carla Hayn
Northwestern University

SYNOPSIS AND INTRODUCTION: Current reporting rules require interperiod tax allocation whereby the income tax expense reported in the income statement is determined on the basis of pretax book (accounting) income, adjusted for permanent differences between the period's taxable income and book income. The temporary differences accumulate on the balance sheet as a deferred tax liability and are assumed to reverse in future years, gradually reducing the liability.

Opponents of interperiod tax allocation argue that reversal of these temporary differences is unlikely or will occur only in the remote future. Thus, support has developed for either partial allocation (e.g., allocating only short-term temporary differences) or no deferral at all. Regardless of such concerns, in 1987 the Financial Accounting Standards Board (FASB) reaffirmed the use of comprehensive interperiod tax allocation.

In this study, we perform cross-sectional analyses relating unexpected stock returns around news disclosures about the Tax Reform Act of 1986 to pertinent firm characteristics in an attempt to assess whether the deferred tax liability is viewed as a liability. The main findings of the study are consistent with the hypothesis that investors view the deferred tax liability as a real liability. They appear to discount it according to the timing and likelihood of the liability's settlement.

The remainder of the article is organized as follows. Section I describes the previous research on the informativeness of tax deferrals,

The financial support of the Accounting Research Center at the J. L. Kellogg Graduate School of Management is gratefully acknowledged. The work has benefitted from comments made by participants in workshops at the University of Arizona, University of California at Berkeley, University of Chicago, University of Illinois, University of Illinois-Chicago Circle, Michigan State University, New York University, Northwestern University, SUNY at Buffalo, Tel-Aviv University, Rutgers Univesity, the University of Washington and the University of Wisconsin.

Submitted September 1989.
Accepted September 1991.

outlines the hypotheses of this study, and details the procedures used to test them. Section II presents and discusses the results, and the final section provides concluding remarks.

Key Words: *Deferred tax liability, Tax deferrals, Interperiod tax allocation, Tax Reform Act of 1986.*

Data Availability: *To the extent allowed by licensing agreements, the authors will make the data used in this study available to others on request upon their completion of a related research project.*

I. Testing the Valuation of the Deferred Tax Liability

THE information content of tax accruals has been examined by several previous research studies.[1] Beaver and Dukes (1972) find that unexpected stock returns are more strongly correlated with earnings measures that incorporate tax deferrals than with earnings measures that do not. They conclude (1972, 331) that "[D]eferral earnings are most consistent with the information set used in setting security prices while cash flow is least consistent."[2] Rayburn (1986), however, reports mixed results on the information content of tax deferrals.

Beaver and Dukes' methodology has some drawbacks: the conclusions are based on a relatively small number of observations in which the highly correlated alternative unexpected earnings measures have opposite signs, proxies were used for the expectation of earnings and its components, and their methodology cannot assess the extent to which the market discounts the deferred tax liability (DTL). A further limitation of their approach is that the economic significance of the obtained association between unexpected returns and reported earnings (which are based on tax deferrals) is unclear. As Beaver and Dukes point out, functional fixation is an alternative explanation of their findings (1972, 331).

The enactment of the Tax Reform Act of 1986 (TRA) provides an opportunity to test the manner by which investors evaluate tax deferrals in a more powerful way. The TRA, among other things, reduced the corporate tax rate from 46 to 34 percent. If the DTL is viewed as a real and imminent liability, this tax rate cut would lead to a proportional decline in the value of this liability and a corresponding increase in the value of the equity.

Accordingly, our study tests the following hypothesis:

H1: The reduction of the corporate income tax rate from 46 to 34 percent by the TRA increases the equity value of firms in proportion to their DTL balance.

The increase in the equity value of an individual firm would depend on the extent to which its DTL is discounted by investors. The discount factor is hypothesized to be a function of the expected growth rate in the DTL and the likelihood of future losses. The higher the expected growth rate, the more distant is the reversal of timing differences and the lower is the liability's present value. Similarly, the higher the likelihood of future losses, the lower is the probability that reversal of temporary differences will result in a tax payment, and, thus, the lower is the expected value of the liability. Our second hypothesis thus addresses the discount factor:

---

[1] Several studies indicate that some accruals have incremental information content (see, e.g., Wilson 1986, 1987; Bowen et al. 1987).

[2] Anthony and Gupta (1988) replicate Beaver and Dukes' study, obtaining similar results.

H2: After controlling for the firm's DTL balance, the reduction in the tax rate increases the firm's equity value by a magnitude that is a function of the likelihood and expected timing of the settlement of its DTL.

To test these hypotheses, we conduct a cross-sectional analysis in which the reduction in the deferred tax liability (ΔDTL), a measure of this liability's growth rate ($GROWTH$), and a measure that captures the probability that the firm will incur a loss for tax purposes ($PROB$) are used to explain each firm's cumulative abnormal return, CAR, over the TRA period. Since the various provisions contained in the TRA are expected to affect the equity value of individual firms differently, we also incorporate pertinent firm characteristics in the analysis. We estimate the following linear regression:

$$CAR_{it} = \alpha_t + \beta_{1t}\Delta DTLM_{it} + \beta_{2t}GROWTH_{it} + \beta_{3t}PROB_{it} + \sum_{j=4}^{n+3}\beta_{jt}CV_{ijt} + \epsilon_{it},$$

where i is the firm index, t is the time index, and each CV_j denotes one of the n control variables. The explanatory variables are discussed below. Computational details for all variables are provided in appendix A.

The main variable of interest, $\Delta DTLM$, is the ratio of the firm's ΔDTL to the market value of its equity.[3] If investors view the DTL as a real obligation, according to hypothesis H1 the estimated coefficient β_1 is expected to be positive. That is, the tax cut, by effectively reducing the market value of the DTL, is expected to increase the value of equity.

The $GROWTH$ variable should capture both the DTL's growth rate and magnitude. $GROWTH_i$ is computed as an interactive variable, $g_i \cdot \Delta DTLM_i$, where g_i is the proxy for expected growth, computed as the compound rate of growth in the ratio of DTL to the book value of the firm's total assets over the most recent five years. Similarly, $PROB$ should capture both the magnitude of the DTL and the probability that the firm will experience a loss for tax purposes. $PROB_i$ is derived as the product $p_i \times \Delta DTLM_i$. The value p_i is the probability that firm i will experience a loss for tax purposes. (The computations of g_i and p_i are detailed in appendix A.)

The change in equity value resulting from the TRA's tax rate reduction is expected to be negatively related to both the firm's DTL growth and its probability of incurring a loss, consistent with hypothesis H2. Thus, we expect both β_2 and β_3 to be negative.

Control Variables

Several firm characteristics have been used in the economics literature to surrogate for the anticipated effect of the TRA on individual firms (Arthur Andersen 1986; Coopers and Lybrand 1986; Dolde 1986; Sinai 1986). In addition, economists have incorporated firm characteristics in theoretical models to predict the effect of the new legislation on capital formation, stock returns, or capital structure.[4]

---

[3] Deflating by the market value of equity allows an intuitive interpretation of the coefficient of the cross-sectional regression (i.e., the change in equity induced by a \$1 decrease in the DTL). It also reduces the potential of a correlated omitted variables problem (see Christie 1987, 237–80).

On the basis of these studies, we incorporate several variables in the cross-sectional regression to control for the effect of the TRA on stock prices. (Computational details are provided in appendix A.) The first control variable, investment tax credit (*ITC*), captures the effect that repeal of this credit had on firm's equity values. We compute this variable by measuring the extent of these benefits that the firm garnered in the past.[5] We expect the coefficient of *ITC* to be negative: the greater the firm's past use of this credit, the greater would be the decline in the market value of its equity as a result of the curtailment of this credit.

The second control variable (*TAX*) proxies for the increase in the firm's tax payment caused by other provisions in the act, such as the alternative minimum tax, designed to equate the effective tax rate across corporations.[6] The change in tax payments resulting from these provisions is expected to be inversely related to the firm's initial effective tax rate. Two alternative measures of the *TAX* variable are used. In the first, the added tax payment is estimated as the difference in the effective tax rate between the period preceding the TRA and the period following it. Given the short post-TRA period,[7] we also use a *TAX* measure based only on the effective tax rate in the pre-TRA period. We expect the coefficient of *TAX* (with either measure) to be negative.

The third control variable is net operating loss carryforwards available for tax purposes (*NOLC*). Post-TRA, companies with carryforwards were in a relatively less favorable position than they were prior to the act for two reasons. First, the decrease in the tax rate reduced the expected value of the tax savings arising from the carryforwards.[8] Second, additional constraints were imposed on the use of such carryforwards, particularly in acquisitions. Because of the resulting decrease in the economic value of the carryforward balances, a negative coefficient on *NOLC* is expected.

In addition to its effect on future tax payments, the repeal of the investment tax credit reduces the net-of-tax return to new equipment while increasing the value of equipment in place.[9] Accordingly, the fourth control variable, *ME*, is the amount of machinery and equipment owned by the firm prior to the implementation of the TRA. This variable may also proxy for the amount of depreciation deductions lost because of the TRA.[10] To the extent that the lost depreciation is not fully captured by the *TAX* variable, the effect of *ME* on equity cannot be stated *a priori*.

---

[4] For example, Auerbach (1987, 1989) and Downs and Hendershott (1987) examine the effect of the TRA on investment levels, economic returns, and equity values. Cutler (1988) and Givoly and Hayn (1991) use firm-specific variables to explain the stock price response of individual firms to the TRA.

[5] The investment tax credit was designed to stimulate capital investment and was computed according to a specified percentage of the cost of qualifying assets.

[6] A primary objective of the TRA was to shift the tax burden to firms that previously enjoyed generous tax breaks. This objective was stated repeatedly by major players in the tax reform legislation. A summary of the act's objective is given by the Joint Committee on Taxation in *General Explanation of the Tax Reform Act of 1986* (U.S. Congress 1987, 8–9).

[7] Data for the post-TRA period are available for only three years: 1987, a transition year when some of the TRA provisions were only partially implemented, and the following two years.

[8] Hayn (1989) finds that these carryforwards affect firm value.

[9] See Auerbach (1987) and Downs and Hendershott (1987) for an estimate of the resulting increase in the economic value of the firm's existing machinery and equipment and in the value of its equity.

[10] Note, however, that the loss in depreciation deductions resulting from the TRA applies only to new assets.

The last control variable, *AGEME*, proxies for the age of the assets in place. For a given level of *ME*, the repeal of the investment tax credit gives firms with newer assets a comparative advantage. The coefficient of this variable is expected to be positive.[11]

Sample and Data

The initial sample used for the study consists of all firms that appear on the 1987 COMPUSTAT Annual Primary/Supplementary/Tertiary Industrial files, with the exclusion of firms that belong to the insurance and banking industries.[12] Return data were retrieved from the CRSP Daily Stock Return and Master files. Data on the amount of *NOLC* available for tax purposes were obtained from the tax footnote in firms' financial statements.[13] Complete return and accounting data were available for 1,348 firms.

Identifying Relevant Test Periods

The pertinent test periods are those in which there is a revision in investors' assessment of the likelihood of the TRA's passage. The ideal test "window" is the one in which the probability assigned by market participants to the passage of the TRA was revised continuously from one extreme (zero probability) to the other (certainty).[14]

The almost constant flow of news reports concerning the TRA over the two years preceding its enactment creates difficulties in identifying the test periods. Given that the TRA legislation was preceded by much public discussion and deliberation, it is apparent, however, that expectations about the TRA were formed gradually over an extended period and, therefore, that the final enactment was partially anticipated.

We searched three major newspapers (the *Washington Post*, *The Wall Street Journal*, and the *New York Times*) and the records of Congressional proceedings concerning the House and Senate versions of the tax reform bill (including the committee and subcommittee hearings) to determine the beginning of the pertinent test period as well as to identify periods during which changes occurred in the public's assessment of the degree of certainty assigned to the likelihood of passage of the TRA. Discussions and proposals pertaining to tax reform appear to have begun in September 1984, which we therefore selected as the beginning of the test period.[15] Subsequently, the pace of events leading to tax reform accelerated rapidly during 1985 and 1986,[16] culminating with the passage of the act by both houses of Congress in September 1986, which marks the end of our test period. The act was signed into law by President Reagan in October 1986.[17]

---

[11] In an extended version of the cross-sectional regression, whose results are not presented, industry dummy variables were added to control for industry effects that might not be captured by the other control variables. Although some industry variables were significant, their inclusion had a negligible effect on the other slope coefficients and on the (adjusted) R^2 of the regression.

[12] These firms (SIC 6022–6200 and 6312–6400) were excluded for two reasons. First, the TRA contained specific provisions pertaining to these industries and controlling for the effect of these provisions would unnecessarily encumber the analysis. Second, many financial institutions have a deferred tax asset rather than a liability, the balance of which is unavailable on COMPUSTAT (since it is combined with other deferred charges).

[13] The tax footnotes for all firms that were likely to have an *NOLC* for tax purposes were examined. These consisted of (1) firms with COMPUSTAT and CRSP data showing a loss in any year from 1983 to 1987, (2) firms that had a NOLC for book purposes in any one of these years, as reported by COMPUSTAT, or (3) firms that had positive earnings in these years but were in the bottom decile of firms with positive earnings in a given year.

[14] Selecting a long enough window may reasonably ensure that, during that period, investors' beliefs were revised over the full probability range. However, by extending the test window, periods during which no revisions occur are added. This is likely to result in an increase in variability ("noise") in the estimation of the parameters of interest and a weakening of the tests.

We identified a total of 130 TRA events within this time interval of 25 months (September 1984 to September 1986). Each news event was characterized as indicating either an increasing or decreasing probability of the passage of the TRA.[18] For each event, we identified an event period consisting of three trading days—the event day and the days preceding and following it. The total number of days contained in the test period is 252.[19]

Although our primary analysis focuses on changes in the firms' equity over the 130 events, we separate the events according to their news content, either increasing or decreasing likelihood of enactment, and then examine the hypotheses over these two subsets of events. This analysis accentuates the relationship between the returns and the independent variables.[20] Further, contrasting the regression results obtained for the two subsets of events provides further evidence on the validity of the events selected; if the events selected are relevant, we would expect the signs of the coefficients to reverse for the two subsets.

To further test the validity of the events identified, as well as for expositional purposes, we classified the events into subperiods. To do this, we first identified five major stages in the TRA's passage. Then, within each stage, we divided the events into those with similar news content regarding the likelihood of the TRA's passage. Thus, each stage is represented by an increasing likelihood and a decreasing likelihood subperiod as shown in table 1 for a total of ten subperiods. To ascertain whether market expectations concerning the passage of the TRA were indeed revised over the identified events, we analyze the returns of individual firms during these subperiods. While different firms might be affected by the TRA to a different extent and in a

[15] During that month, Representative Rostenkowski, chairman of the House Ways and Means Committee, endorsed tax reform and made it a major item on the House Ways and Means Committee agenda. President Reagan also expressed his support for tax reform and Treasury Secretary Regan introduced the Treasury's first attempt at a tax reform plan in October 1984.

[16] The upsurge in public interest in revising the tax code is evident from the extent of news coverage on the issue. For example, the number of news items relating to tax overhaul in the *Washington Post* was 11 in 1982, 37 in 1983, 30 in the first nine months of 1984, 84 in the last three months of 1984, 515 in 1985, and 506 in 1986. The two other newspapers examined also reflect this dramatic increase in the number of articles on tax revision beginning in the last quarter of 1984.

[17] The passage of the act by both houses of Congress is selected as the end of the test period because, at this time, enactment became certain. President Reagan had been a strong supporter of tax reform over the preceding two years, and his approval of the act was a foregone conclusion.

[18] Throughout the period examined, there was a great deal of uncertainty regarding the eventual passage of this far-reaching tax reform, with many skeptics expressing serious doubts about its viability only a few days before the actual passage of the act.

[19] Several events occur on consecutive days. Usually, the content of the news occurring on these consecutive days was the same (i.e., all indicating either an increasing or decreasing likelihood of the act's passage). For these events, we included the days on either side of these events as well as the event days in the test period. In a few cases, the events occurring on consecutive days were of opposite news content (i.e., one event indicating an increasing likelihood of enactment, while the other indicated a decreasing likelihood). We excluded these events from the analysis.

[20] If the selection of the subperiods is correct, the probability of the TRA's passage during the "increasing likelihood" events must have cumulatively increased by more than 1.0.

Table 1
Description of Events Leading to Enactment of the Tax Reform Act of 1986 (TRA)

Major Stages of TRA Legislation	*Major Events*
I. Introduction of "Treasury I"	
Subperiod 1: News events indicating increasing likelihood of TRA's passage.	The White House introduced "Treasury I," the first tax reform plan. House Democrats, led by Rostenkowski, chairman of the House Ways and Means Committee, strongly support tax reform.
Subperiod 2: News events indicating decreasing likelihood of TRA's passage.	Congress strongly opposed "Treasury I." The White House was delinquent in introducing a revised tax reform plan, leading many to doubt that President Reagan was serious about tax reform.
II. Introduction of "Treasury II"	
Subperiod 3: News events indicating increasing likelihood of TRA's passage.	A new, compromise tax plan was presented to Congress. President Reagan began campaigning heavily for congressional support.
Subperiod 4: News events indicating decreasing likelihood of TRA's passage.	Congressional and public support for tax reform waned due to lack of enthusiasm for new plan. President Reagan decided to shift his attention elsewhere.
III. Passage of the House Bill	
Subperiod 5: News events indicating increasing likelihood of TRA's passage.	The House Ways and Means Committee drafted a tax reform bill. It was approved by the House after intensive lobbying by House Democrats and President Reagan.
Subperiod 6: News events indicating decreasing likelihood of TRA's passage.	Congressional and public support for tax revision slipped further. Serious disputes arose between the White House and the House Ways and Means Committee. House Republicans abandoned President Reagan in his drive for tax reform.
IV. Passage of the Senate Bill	
Subperiod 7: News events indicating increasing likelihood of TRA's passage.	New tax plan that was considered more lenient was offered by the Senate Finance Committee, led by Packwood. Compromises reached with the House and the White House. The committee passed a bill that was subsequently approved by the full Senate.
Subperiod 8: News events indicating decreasing likelihood of TRA's passage.	The House and the White House rejected the Senate Finance Committee's tax bill. Some felt that the many "tax breaks" included would kill tax reform.
V. Passage of the Tax Reform Act of 1986	
Subperiod 9: News events indicating increasing likelihood of TRA's passage.	Conference members from the House and Senate, appointed to draft a bill for submission to Congress, reached agreement on a number of key points. They drafted a bill that gained strong presidential support. Both the House and the Senate passed the bill.
Subperiod 10: News events indicating decreasing likelihood of TRA's passage.	Members of the tax conference argued over various provisions of the final bill.

different direction, we expect a given firm's returns in any two subperiods to be positively correlated if the news in the two subperiods is of a similar content (e.g., increasing likelihood in both subperiods) and negatively correlated if the news in the two subperiods is of an opposite content.[21] Accordingly, we compute the cross-sectional correlations of firms' returns between all pairs of the ten subperiods.[22]

The correlation results are presented in table 2. Thirty-seven of the 45 correlations are in a direction consistent with the assumption that the TRA-related news in the subperiods affected firms' returns in the assumed direction: positive (negative) correlations are obtained for most pairs with similar (opposite) news content.[23]

These results do not necessarily imply that the identified test period contains all pertinent events nor do they mean that irrelevant events are not included. They do suggest, however, that the selected events are relevant for our tests.

II. Results

Descriptive statistics for the variables of the 1,348 sample firms are presented in table 3. The mean expected reduction in DTL (scaled by the market value of equity, $\Delta DTLM$) is 2.9 percent. The exact amount of increase in a given firm's equity resulting from the reduction in this liability depends on the extent to which it is discounted.

There is, on average, an increase in the relative weight of the DTL over the preceding five years. This is reflected in the compound annual growth measure, g_i, which has a mean value of 6 percent, with 25 percent of the firms showing a growth rate of more than 12.8 percent. The probability of reporting a loss for tax purposes in a given year, p_i, is 21 percent across the sample of firm years. However, most firms did not experience a loss in the period examined (the median value of p_i is zero).

The value of the *ITC* variable, representing the annual amount of investment tax credit lost as a result of the TRA, is, on average, 1.2 percent of the market value of the firm. About 20 percent of the firms had *NOLC* available for tax purposes, which amounts to an average of 15.4 percent of the market value of the sample. The average effective

---

[21] Note that the content of the proposed act itself evolved over the period examined. The stock price response to the TRA-related news is expected, therefore, to reflect not only the likelihood of the act's passage, but also its expected content. However, the primary objectives of the tax reform (e.g., a more equitable sharing of the tax burden across firms) and thus its main proposed provisions (a tax rate cut coupled with the elimination of tax preferences) were not altered over the period. Therefore, cross-sectional differences in the *direction* of stock price movements in response to TRA-related news could be viewed as stemming primarily from the changing likelihood of the act's passage.

[22] Noreen and Sepe (1981) use a similar approach for assessing bi-directional effects. The correlations are between raw returns. The correlations between abnormal returns will not capture new aggregate effects, if any, of the TRA on the stock market.

[23] To rule out the possibility of an omitted variable that might induce serial correlation in the returns of firms and thus invalidate the conventional significance test, we generated "simulation statistics" which serve as the null parameters in testing the hypothesis that the sample correlation coefficients are no different from those prevailing in "normal" times. To generate these statistics, we replicate our test period using randomly selected time intervals of the same length (about 25 months) and that mirror the subperiods in their event composition. We then compute the correlation in firms' returns between any pair of the replicated subperiods, producing 45 ($9 \times 10/2$) pairwise correlation coefficients. We repeat this procedure 100 times, each time randomly replicating our test period in the manner described. At the end of the process, we obtain the "simulation statistics" (e.g., the mean and standard deviation of the correlation coefficient for each of the 45 combinations of paired subperiods). The means produced by the simulation are very close to zero (ranging from -0.021 and 0.015). The refinement provided by the simulation-based statistics leads to essentially identical inferences.

Table 2
Correlation Coefficients between Firms' Returns in Different Subperiods of TRA-related News Events

(n=1,348)

	Subperiods[a]										Combined Subperiods[b]	
Subperiod	1	2	3	4	5	6	7	8	9	10	Increasing Likelihood	Decreasing Likelihood
1	1.00	−0.36*	0.14*	−0.05*	0.08*	−0.03	0.13*	−0.04**	0.17*	−0.06*	0.22*	−0.09*
2		1.00	−0.09*	0.04	−0.11*	0.07*	−0.06*	0.05*	−0.04	−0.08*	−0.13*	0.04**
3			1.00	−0.09*	0.08*	0.03	0.08*	0.02	0.06*	−0.03	0.14*	−0.05*
4				1.00	−0.04	0.06*	−0.03	0.03	−0.02	−0.09*	−0.06*	0.03
5					1.00	−0.11*	0.08*	−0.07*	0.00	0.06*	0.09*	−0.08*
6						1.00	0.04**	0.09*	−0.12*	−0.03	−0.08*	0.09*
7							1.00	−0.17*	0.05*	−0.10*	0.14*	−0.14*
8								1.00	−0.08*	−0.03	−0.11*	0.10*
9									1.00	−0.14*	0.12*	−0.13*
10										1.00	−0.12*	−0.08*
Combined Subperiods												
Increasing likelihood of enactment											1.00	−0.25*
Decreasing likelihood of enactment												1.00

* Significant at a level of 5 percent.

** Significant at a level of 10 percent.

[a] See table 1 for identification of the subperiods.

[b] In computing the correlation between returns in a particular subperiod and the combined subperiods, the particular subperiod return is excluded from the calculation of the returns for the combined subperiods.

Table 3
Distribution Statistics for Variables in the Cross-sectional Regression
($n=1{,}348$)

	*Variables**								
	CAR	*ΔDTLM*	*g*	*p*	*ITC*	*TAX*	*NOLC*	*ME*	*AGE*
Mean	−0.011	0.029	0.060	0.210	0.012	0.005	0.164	0.445	0.511
Standard Dev.	0.212	0.043	0.203	0.305	0.040	0.047	0.632	0.719	0.440
First Quartile	−0.142	0.002	0.000	0.000	0.000	−0.008	0.000	0.134	0.310
Median	−0.014	0.013	0.011	0.000	0.003	0.000	0.000	0.281	0.452
Third Quartile	0.122	0.035	0.128	0.348	0.013	0.012	0.000	0.506	0.647

* Detailed definitions and computations are provided in the text and appendix A.

CAR = cumulative abnormal return for the entire TRA period;
ΔDTLM = reduction in deferred tax liability (DTL) due to corporate tax rate cut, deflated by market value of equity;
g = annual growth rate in the ratio of the DTL to the book value of assets (used in computing the *GROWTH* variable);
p = probability of a loss for tax purposes (used in computing the *PROB* variable);
ITC = average annual amount of investment tax credit eliminated by TRA, deflated by market value of equity;
TAX = added amount of tax payments due to change in effective tax rate;
NOLC = net operating loss carryforward, deflated by market value of equity;
ME = amount of machinery and equipment in place, deflated by market value of equity; and
AGE = age of machinery and equipment in place (used in computing the *AGEME* variable).

tax rate is 39.4 percent in the pre-TRA period, which is significantly lower than the statutory rate of 46 percent.

The main results of estimating the cross-sectional regression are shown in table 4. The results are reported for the entire period (consisting of the 130 events) and separately for the two subsets containing news events of similar content.

The most relevant test period is the entire period, since it is over this period that the probability investors assigned to the eventual passage of the TRA is hypothesized to have been revised through the entire probability range from zero to 1. The coefficients resulting from this analysis for all of the variables are significant at the 5 percent level and have the expected signs. Furthermore, consistent with the hypothesized effect of the TRA, all coefficients reverse signs between the increasing and decreasing likelihood subsets of events.

The main variable of interest, *ΔDTLM*, has a positive and significant coefficient for the regression over the entire period, indicating that the anticipated decline in the nominal value of the deferred tax liability leads to an equity appreciation. The negative coefficients on *GROWTH* and *PROB* in this period suggest that this appreciation varies cross-sectionally with the expected timing and likelihood of the DTL's settlement. Specifically, for a given DTL balance, firms whose DTL balances have grown faster and firms with a higher probability of reporting a tax loss in any given year experienced lower equity appreciation than other firms.

Table 4
Results for the Cross-sectional Regression

$$CAR_{it}=\alpha_t+\beta_{1t}\Delta DTLM_{it}+\beta_{2t}GROWTH_{it}+\beta_{3t}PROB_{it}+\beta_{4t}ITC_{it}+\beta_{5t}TAX_{it}+\beta_{6t}NOLC_{it}+\beta_{7t}ME_{it}+\beta_{8t}AGEME_{it}+\epsilon_{it}$$

($n=1{,}348$; t-values are given in parentheses)

Entire Period

Intercept	*ΔDTLM*	*GROWTH*	*PROB*	*ITC*	*TAX*	*NOLC*	*ME*	*AGEME*	*Adjusted* R^2 *(F)*
−0.033 (−4.01)*	0.558 (3.30)*	−0.184 (−2.67)*	−0.386 (−3.28)*	−0.636 (−3.66)*	−0.595 (−4.65)*	−0.012 (−2.33)*	0.019 (2.20)*	0.069 (3.20)*	0.068 (12.96)*

"Increasing Likelihood" Events[a]

Intercept	*ΔDTLM*	*GROWTH*	*PROB*	*ITC*	*TAX*	*NOLC*	*ME*	*AGEME*	*Adjusted* R^2 *(F)*
−0.058 (−8.32)*	0.745 (3.71)*	−0.226 (−2.83)*	−0.341 (−4.98)*	−0.876 (−4.71)*	−0.734 (−5.30)*	−0.023 (−2.81)*	0.036 (3.73)*	0.082 (3.54)*	0.106 (20.44)*

"Decreasing Likelihood" Events[a]

Intercept	*ΔDTLM*	*GROWTH*	*PROB*	*ITC*	*TAX*	*NOLC*	*ME*	*AGEME*	*Adjusted* R^2 *(F)*
0.025 (9.35)*	−0.187 (−2.80)*	0.042 (1.88)*	0.145 (1.58)**	0.240 (2.50)*	0.139 (3.78)*	0.011 (3.26)*	−0.017 (−3.88)*	− 0.013 (−3.03)*	0.041 (8.01)*

Note: Detailed definitions and computations are provided in the text and appendix A.

CAR = cumulative abnormal return for the entire TRA period (or for the subset of events);
ΔDTLM = reduction in DTL due to corporate tax rate cut, deflated by market value of equity;
GROWTH = an interactive variable equal to g • *ΔDTLM*, where g is the compound growth rate in the ratio of the firm's DTL to the book value of its assets;
PROB = an interactive variable equal to p • *ΔDTLM*, where p is the probability that the firm will experience a loss for tax purposes;
ITC = average annual amount of investment tax credit eliminated by TRA, deflated by market value of equity;
TAX = proxy for added amount of tax payments due to increase in effective tax rate;
NOLC = amount of net operating loss carryforward, deflated by market value of equity;
ME = amount of machinery and equipment in place, deflated by market value of equity; and
AGEME = an interactive dummy variable that receives the value of *ME* when age of machinery and equipment is below average, 0 otherwise.

* Significant at the 5 percent significance level.

** Significant at the 10 percent significance level.

[a] "Increasing likelihood" and "decreasing likelihood" events are those in which the TRA-related news reflects an increase or decrease, respectively, in the likelihood of the TRA's passage.

Further support for this conclusion is provided by analyzing the two subsets of similar-news-content events. The signs of the coefficients of *ΔDTLM*, *PROB*, and *GROWTH* for the increasing likelihood events are the same as those for the entire period, whereas they reverse for the decreasing likelihood events. The reversal of the signs of the regression coefficients provides further evidence that our identification of the content of the events is valid.

The control variables are all significant at the 5 percent level and have the expected directions. Note that the coefficient of *ME* is positive, which suggests that firms with assets in place have a comparative advantage over others due to the repeal of the investment tax credit on new acquisitions.

Further insight on how investors discount the DTL can be gained from examining the magnitude of the regression coefficients. The coefficient of $\Delta DTLM$ for the entire period is 0.558, suggesting that the DTL is being discounted by investors. Conditional on the control variables, one dollar of DTL is valued, on average, at about 56 cents. For a firm that belongs to the fourth quartile of the population in terms of growth of its DTL balance (i.e., a firm whose g_i is equal to 0.128, as shown in table 3) and probability of a loss (i.e., p_i is equal to 0.348), one dollar of DTL appears to be more heavily discounted, valued at only 40 cents.[24]

If discounting of the DTL takes place, we would also expect the discount factor to be correlated with firm risk. That is, the magnitude of the $\Delta DTLM$ coefficient for a given subsample of firms would decrease with the risk of the firms in that subsample. Accordingly, we partitioned the sample into two: firms with above median betas and firms whose betas were equal to or below the median beta. The estimates of the $\Delta DTLM$ coefficients for the two groups for the entire test period are 0.82 and 0.39, respectively (both significant at the 5 percent level), consistent with the economic interpretation of the coefficient as a discount factor.

Diagnostic Checks

Several statistical problems may impair the inferences drawn from the results reported in table 4. One is misspecification of the market model (e.g., the selected variables may proxy for omitted variables). In this case, the estimated regression coefficients could differ from zero without implying a rejection of the respective null hypotheses.

The presence of cross-sectional dependence in the regression residuals may also lead to incorrect inferences. It has been documented that such dependence, resulting possibly from dependence in contemporaneous abnormal returns of different firms, is minor for daily or weekly abnormal returns but becomes stronger for longer accumulation periods.[25] Cross-sectional dependence results in a downward bias in the estimates of the regression coefficients' variance.[26] Other potential statistical problems are multicollinearity and heteroskedasticity, which induce inefficiency in estimation.

Diagnostic tests and corrective procedures are presented in appendix B. They provide evidence that our previously mentioned inferences are not affected by the statistical problems discussed above. In particular, (1) no spurious relationship appears to exist between the dependent and explanatory variables, (2) cross-sectional dependence in residuals is not influential, and (3) no serious multicollinearity in the data or heteroskedasticity of the disturbance term of the regression appear to be present.

---

[24] This is computed by using the value of the coefficients as follows:

$$(0.558-[0.184\times 0.128]-[0.386\times 0.348]).$$

[25] See Collins and Dent (1984) and Bernard (1987).

[26] See Christie (1987) and Bernard (1987) for further discussion of cross-sectional dependence.

III. Concluding Remarks

This study shows that investors view deferred taxes as a real liability and that they discount it to its present value according to the likelihood and timing of its settlement. For users of financial statements, our findings suggest that a portion of the deferred tax liability should be better viewed as part of equity. For accounting rule-making bodies, the results indicate that deferred taxes, arising from comprehensive interperiod tax allocation, are being transformed by investors into a value that appears to be consistent with the notion of partial allocation.

The FASB recently suspended FAS 96 and issued a new exposure draft on accounting for income taxes. Both documents maintain the principle of comprehensive allocation. Our results, however, lend support to changing the principle in favor of selective, partial allocation.

Appendix A
Definitions of Variables in the Cross-sectional Regression

Variable and Definition	*Computation*
CAR Cumulative abnormal returns	Cumulative abnormal returns are computed by compounding abnormal returns over the test period, t. Abnormal returns are calculated as $R_{id}-E(R_{id})$. R_{id} is the actual return and $E(R_{id})$ is the expected return of firm *i* on day *d*. $E(R_{id})$ is assumed to be generated by the market model $E(R_{id})=\hat{\alpha}_i+\hat{\beta}_i R_m$ using daily stock returns over the 75 trading days preceding and the 75 trading days following the test period. In estimating the market model, R_m is proxied by the value-weighted return of all firms on the NYSE and AMEX. The Scholes and Williams (1977) procedure that reduces the problems associated with nonsynchronous trading is used in the estimation.
ΔDTLM Reduction in the deferred tax liability	Computed as the balance of the DTL account at year-end times (0.46−0.34)/0.46, divided by the market value of the firm's equity at the year-end.
GROWTH Growth in the ratio of the DTL to the book value of assets	Computed as *ΔDTLM* • g. The value of g is computed as the geometric mean annual growth in the ratio of the DTL to the book value of the assets (*BVA*) over the recent five years, as follows: $g=[(DTL_\tau/BVA_\tau)/(DTL_{\tau-5}/BVA_{\tau-5})]^{1/5}-1,$ where τ is the year index.
PROB Probability of not paying taxes	Computed as *ΔDTLM* • *p*. The probability that a firm will not pay taxes in a given future year, p, is based on the number of years in the most recent five-year period in which the firm did not have to pay taxes (before applying the investment tax credit (*ITC*)). Because recent experience is likely to be more relevant for assessing this probability, more recent years were given a greater weight (a variable based on an equal weighting scheme did not change the regression results materially). Each year τ was assigned a value, p_τ, where p_τ equals 0 if the firm paid taxes in year τ (i.e., the amount of current taxes, before applying the *ITC*, is positive) and 1 if the firm did not pay taxes in year τ. Geometrically declining weights (by 35 percent each year), w_τ, were assigned to the five years; the most recent year had a weight of 0.40 and the fifth year had a weight of 0.07. (Use of a ten-year estimation period produced similar results.) P_t is defined as the weighted average of p_τ: $p_t=\sum_{\tau=t-5}^{t} p_\tau \bullet w_\tau.$

Appendix A—*Continued*
Definitions of Variables in the Cross-sectional Regression

Variable and Definition	*Computation*
ITC Past use of investment tax credits	Computed as the average investment tax credits used over the most recent five-year period divided by the market value of the firm's equity at the end of 1986.
TAX Proxy for the added tax payment due to the increase in effective tax rate	Computed as the difference between the firm's effective tax rate (ETR) in the post-TRA period (1987–89) and the ETR over the pre-TRA period (1981–86), times the average pretax income over the pre-TRA period, deflated by the market value of the firm's equity at the end of 1986. The firm's effective tax rate is defined as the ratio of the total current income taxes (before the investment tax credit) divided by pretax income where the numerator and the denominator each represent an aggregation over the relevant years. (The *TAX* variable was also computed by using the statutory tax rate in the post-TRA period (0.34) rather than the effective tax rate. The results with this measure are essentially the same as those for the above definition of *TAX*, and therefore only the latter are reported.) An alternate measure of *TAX*, *TAX*1 is also used. This measure is computed as the average ETR over the ten years preceding the TRA period.
NOLC Net operating loss carryforward	Computed as the ratio of the net operating loss carryforward available for tax purposes at the end of 1986 deflated by the market value of the firm's equity at the end of 1986.
ME Machinery and equipment in place	Computed as the ratio of machinery and equipment on the firm's books at the end of 1986 divided by the market value of the firm's equity at the end of 1986.
AGEME Age of machinery and equipment in place	Assigned a value of 0 if AGE_i is less than or equal to the average AGE_i of the sample, and a value of ME_i if AGE_i is greater than the average AGE_i of the sample. AGE_i is computed as the sum of capital expenditures in 1985 and 1986, divided by net plant, property, and equipment at the end of 1986.

Appendix B
Diagnostic Tests and Corrective Procedures

Misspecification

To examine whether our inferences are due to misspecification of the model, the cross-sectional regression was estimated for 100 essentially non-overlapping, non-event periods.[27] Each non-event period had the same chronological length as the test period (about 25 months) and contained subperiods of the same length, chronological order, and separating intervals as the original test subperiods.[28] The mean *t*-values of each of the regression coefficients in the non-event periods are close to zero. The null hypothesis that an individual variable's *t*-value is distributed normally with a zero mean and a standard deviation of 1 (using the χ^2 test) could not be rejected for all variables.

---

[27] Similar randomization procedures were used by Leftwich (1981) and Lys (1984).

[28] A slight overlapping (of about 5 percent of the days) was necessary for the construction of 100 such periods from the available CRSP data.

Table B1

The t-values of Cross-sectional Regression Coefficients Corrected for Bias from Cross-sectional Dependence

$$CAR_{it}=\alpha_t+\beta_{1t}\Delta DTLM_{it}+\beta_{2t}GROWTH_{it}+\beta_{3t}PROB_{it}+\beta_{4t}ITC_{it}+\beta_{5t}TAX_{it}$$
$$+\beta_{6t}NOLC_{it}+\beta_{7t}ME_{it}+\beta_{8t}AGEME_{it}+\epsilon_{it}$$

Period Examined	*Intercept*	*ΔDTLM*	*GROWTH*	*PROB*	*ITC*	*TAX*	*NOLC*	*ME*	*AGEME*
Entire Period	−1.21	2.18*	−1.36	−2.52*	−2.14*	−2.29*	−1.70**	1.75**	2.85*
"Increasing Likelihood" Events•	−3.12*	2.34*	−1.89*	−2.25*	−2.52*	−3.03*	−2.31*	2.45*	3.65*
"Decreasing Likelihood" Events•	4.61*	−1.92*	1.53	1.26	2.05*	3.12*	2.02*	−1.85**	−2.34*

Note: Detailed definitions and computations are provided in the text and appendix A.

CAR=cumulative abnormal return for the entire TRA period (or for the subset of events);
ΔDTLM=reduction in DTL due to corporate tax rate cut, deflated by market value of equity;
GROWTH=an interactive variable equal to g•*ΔDTLM*, where g is the compound growth rate in the ratio of the firm's DTL to the book value of its assets;
PROB=an interactive variable equal to p•*ΔDTLM*, where p is the probability that the firm will experience a loss for tax purposes;
ITC=average annual amount of investment tax credit eliminated by TRA, deflated by market value of equity;
TAX=proxy for added amount of tax payments due to increase in effective tax rate;
NOLC=amount of net operating loss carryforward, deflated by market value of equity;
ME=amount of machinery and equipment in place, deflated by market value of equity; and
AGEME=an interactive dummy variable that receives the value of *ME* when age of machinery and equipment is below average, 0 otherwise.

* Significant at the 5 percent significance level.
** Significant at the 10 percent significance level.
• "Increasing likelihood" and "decreasing likelihood" events are those in which the TRA-related news reflects an increase or decrease, respectively, in the likelihood of the act's passage.

Cross-sectional Dependence in the Residuals

We use the procedure suggested by Bernard (1987) to correct for the bias in the estimated standard error of the regression coefficients caused by potential cross-sectional dependence in the residuals. Under this procedure, we calculate for each coefficient β_k the ratio R_k between its time-series variance ($\sigma^2(B_k)$, based on the values of the coefficients derived from a series of cross-sectional OLS regressions) to the mean (over a series of cross-sectional OLS regressions) of its regression variance.[29] A ratio of 1 suggests no bias. A value greater than 1 means that the OLS estimate of the coefficient's variance understates the true variance. The above ratio is used to "correct" the reported results for potential bias.

---

[29] In other words,

$$R_k=\frac{1/(N-1)\sum_{n=1}^{N}(\beta_{kn}-\bar{\beta}_k)^2}{1/N\sum_{N=1}^{N}\mathrm{Var}(\hat{\beta}_{kn})}\qquad n=1,\ldots,N,$$

where:

$$\bar{\beta}_k=\frac{1}{N}\sum_{n=1}^{N}\beta_{kn},$$

n is the regression index, and *N* is the total number of cross-sectional regressions.

To obtain a time-series of β_k and of $\sigma^2(\beta_k)$, we estimate the cross-sectional regression over the 100 non-event periods. Most of the values of R_k generated by this procedure are between 1.5 and 3.6, indicating a downward bias in the estimates of the coefficients' variance.[30] The corrected *t*-values for the combined period and the "increasing likelihood" and "decreasing likelihood" subperiods are provided in table B1. Although the new *t*-values are lower, as expected, the coefficient of the variable of most interest, $\Delta DTLM$, maintains a significance level of 5 percent in all of the subperiods examined.

Multicollinearity

The results do not seem to be affected by multicollinearity. The corrrelation coefficients between the explanatory variables are mostly between −0.10 and 0.10. The only high correlations are between $\Delta DTLM$ and its two related interactive variables, *GROWTH* and *PROB* (0.32 and 0.34, respectively). This dependence, however, is induced by the manner in which these variables are constructed.

Heteroskedasticity

To correct for this problem, we reestimated the cross-sectional regression with each individual observation deflated by the standard deviation of the firm's abnormal return. The latter is estimated from the 100 non-event periods described earlier. The results are similar to those obtained without the correction.

---

[30] These values appear to be somewhat higher than those derived by Bernard (1987) for two typical cross-sectional studies of abnormal return behavior. The higher values may be due to our larger sample: in the absence of a proportional increase in the number of industries represented in the sample, an increase in the sample size adds to the bias.

References

Anthony, J. H., and S. Gupta. 1988. Empirical tests of investor usefulness for tax basis financial statements. Working paper, Michigan State University, East Lansing.

Arthur Andersen & Co. 1986. *Tax Reform 1986: Analysis and Planning.* Arthur Andersen & Co.

Auerbach, A. J. 1987. The Tax Reform Act of 1986 and the cost of capital. *Journal of Economic Perspectives* 1 (Summer): 73–86.

———. 1989. Tax reform and adjusted costs: The impact on investment and market value. *International Economic Review* 30 (November): 936–62.

Beaver, W. H., and R. E. Dukes. 1972. Interperiod tax allocation, earnings expectations and the behavior of security prices. *The Accounting Review* 47 (April): 320–32.

Bernard, V. L. 1987. Cross-sectional dependence and problems in inference in market-based accounting research. *Journal of Accounting Research* 25 (Spring): 1–48.

Bowen, R. M., D. Burgstahler, and L. A. Daley. 1987. The incremental information content of accrual versus cash flows. *The Accounting Review* 62 (October): 723–47.

Christie, A. A. 1987. On cross-sectional analysis in accounting research. *Journal of Accounting & Economics* 9 (March): 231–58.

Collins, D. W., and W. T. Dent. 1984. A comparison of alternative testing methodologies used in capital market research. *Journal of Accounting Research* 22 (Spring): 48–84.

Coopers & Lybrand. 1986. *Tax Reform Act of 1986.* Washington, D.C.: Coopers & Lybrand.

Cutler, D. M. 1988. Tax reform and the stock market: An asset price approach. *The American Economic Review* 78 (December): 1107–17.

Dolde, W. 1986. The impact of corporate tax reform by S & P industries. *Shearson-Lehman Brothers Prospects* (May 23). Shearson-Lehman Brothers.

Downs, T., and P. H. Hendershott. 1987. Tax policy and stock prices. *National Tax Journal* 40 (June): 183–90.

Givoly, D., and C. Hayn. 1991. The aggregate and distributional effect of the Tax Reform Act of 1986 on firm valuation. *Journal of Business* 64 (July): 363–92.

Hayn, C. 1989. Tax attributes as determinants of shareholder gains in corporate acquisitions. *Journal of Financial Economics* 23 (April): 121–54.

Leftwich, R. 1981. Evidence on the impact of mandatory changes in accounting principles on corporate loan agreements. *Journal of Accounting & Economics* 3 (March): 3–36.

Lys, T. 1984. Mandated accounting changes and debt covenants: The case of oil and gas accounting. *Journal of Accounting & Economics* 6 (December): 39–65.

Noreen, E., and J. Sepe. 1981. Market reactions to accounting policy deliberations: The inflation accounting case. *The Accounting Review* 56 (April): 253–69.

Rayburn, J. 1986. The association of operating cash flows and accruals with security returns. *Journal of Accounting Research* 24 (Supplement): 112–33.

Scholes, M., and J. Williams. 1977. Estimating betas from nonsynchronous data. *Journal of Financial Economics* 5: 309–28.

Sinai, A. 1986. The Tax Reform Act of 1986: Winners and losers. *Shearson-Lehman Brothers Economic Studies Series*. Report 22. Shearson-Lehman Brothers.

U.S. Congress, Joint Committee on Taxation. 1987. *General Explanation of the Tax Reform Act of 1986*. Washington, D.C.: U.S. Government Printing Office.

Wilson, G. P. 1986. The relative information content of accrual and cash flows: Combined evidence at the earnings announcement and annual release date. *Journal of Accounting Research* 24 (Supplement): 165–200.

———. 1987. The incremental information content of accrual and funds components of earnings after controlling for earnings. *The Accounting Review* 62 (April): 293–322.

THE ACCOUNTING REVIEW
Vol. 67, No. 2
April 1992
pp. 411–426

The Role of the Accounting Rate of Return in Financial Statement Analysis

Richard P. Brief
New York University
Raef A. Lawson
State University of New York at Albany

SYNOPSIS AND INTRODUCTION: The accounting rate of return (*ARR*) is "not only a central feature of any basic text on financial statement analysis but also figures commonly in the evaluation by investment analysts of the financial performance of firms" (Whittington 1988, 261). Notwithstanding the prominence given to this financial ratio (Foster 1986, 77–79), many writers have warned that the *ARR* lacks economic significance and can be a very misleading measure of profitability. Fisher and McGowan (1983, 90), for example, conclude that "there is no way in which one can look at accounting rates of return and infer anything about relative economic profitability." In the same vein, Rappaport (1986, 31) states flatly that the comparison of the *ARR* with the cost of capital is "clearly like comparing apples with oranges."

The limitations of using *ARR*s to estimate the economic rate of return have been discussed over the last 25 years; for example, Harcourt (1965), Solomon (1966), Kay (1976), Fisher and McGowan (1983), Salamon (1985), Edwards et al. (1987), and Brief and Lawson (1991). This research has focused mainly on the question of whether the *ARR* is a good proxy for the economic return and the literature contains virtually no discussion of other ways in which the *ARR* might be used in financial analysis. An important exception is Peasnell (1982b) who presents a common analytical framework connecting conventional economic concepts of value and yield and accounting models of profit and return. However, even here, the main emphasis is on the relationship between accounting and economic rates of return. This emphasis is quite evident in Peasnell's concluding comment

The authors thank Joshua Livnat, Andrew Stark, and the reviewers for their helpful comments and suggestions.

Submitted January 1991.
Revised September 1991.
Accepted November 1991.

that "it is difficult to assign economic significance to accounting yields except either (1) as surrogate measures of IRR or (2) when they are defined in terms of entry- or exit-market prices" (379–80).

A different slant on the economic significance of *ARR*s is presented here where the focus is on the use of the *ARR* in the valuation process, not in the determination of profitability. The purpose is to show that an expression for an accounting-based measure of discounted cash flows (DCF) can be derived in terms of the *ARR*. Thus, quite apart from the question of whether or not the *ARR* is an accurate estimate of the economic rate of return, this financial ratio has a key role to play in the valuation process.

The results have both analytical and practical significance. On an analytical level, the derivations can be viewed as basic "bookkeeping relationships" that are associated with a double-entry system. On the more practical side, while the DCF techniques currently used in practice "are largely *accounting*-based valuation approaches" (DeAngelo 1990, 100), the use of accounting data in DCF valuations is very indirect. Present practice is to "use historical accounting relationships to forecast future earnings, from which future cash flows are estimated" and, in addition, to estimate terminal values "from projections of future earnings" (p. 100). However, instead of basing DCF valuations on cash flows which are derived from accounting data, a more direct method of analysis is to base the DCF valuations directly on accounting data. Understanding how accounting data can be used in *DCF* analysis leads to a greater appreciation of the general nature of accounting and provides a compelling reason to give the *ARR* a more prominent place in financial statement analysis.

Key Words: *Accounting rate of return, Discounted cash flow analysis, Double-entry bookkeeping, Valuation.*

THE present value of a project (which could be a division of a firm or the firm itself) is the discounted value of free cash flows out to a *valuation horizon*, plus the forecasted value of the project at the horizon, discounted back to present value (Brealey and Myers 1988, 63). The first section of this paper derives a general accounting-based formula for the present value of a project. The second section simplifies the formula in two steps. First, and with no loss of generality, a constant *ARR*, called the *pseudo* internal rate of return (*IRR*), can be substituted for the variable, single-period rates of return. Second, a constant-growth model is assumed. Several other special cases also are derived. The third section considers the use and practical limitations of the DCF formula and deals with questions relating to the calculation of the constant *ARR* as well as the error introduced by the assumption of constant growth. An example illustrating the application of the accounting-based DCF model also is presented.

I. Derivation of Accounting-based DCF Formula

Let V_0 be the present value of the project at the end of period 0, C_t the net cash flows in period t, R_n the residual value of the project, and k the cost of capital.[1] The time horizon is any time segment during the life of the project where $t=0$ is the beginning of period 1 and $t=n$ is the last period in the time slice. If the project is initiated in $t=0$ and terminates in $t=n$, the time segment is the life of the project. The standard definition of V_0 is:

$$V_0=\sum_{t=1}^{n}\frac{C_t}{(1+k)^t}+\frac{R_n}{(1+k)^n}. \tag{1}$$

This section derives the accounting-based equivalent of V_0.

Define I_t as accounting income in period t and A_t as accounting book value at the end of period t and assume that accounting income is based on a comprehensive income concept. Also, define the *ARR* in period t, a_t, as:

$$a_t=I_t/A_{t-1}.$$

From the accounting identity:

$$I_t=C_t+(A_t-A_{t-1}),$$

it immediately follows that:

$$C_t=(1+a_t)A_{t-1}-A_t.$$

The accounting-based equivalent of:

$$\sum_{t=1}^{n}\frac{C_t}{(1+k)^t}$$

is derived by replacing C_t with $(1+a_t)A_{t-1}-A_t$. This substitution links traditional DCF analysis with accounting information. Thus:

$$\sum_{t=1}^{n}\frac{C_t}{(1+k)^t}=\sum_{t=1}^{n}\frac{(1+a_t)A_{t-1}}{(1+k)^t}-\sum_{t=1}^{n}\frac{A_t}{(1+k)^t}.$$

Since:

$$\sum_{t=1}^{n}\frac{A_t}{(1+k)^t}=(1+k)\sum_{t=0}^{n-1}\frac{A_t}{(1+k)^{t+1}}+\frac{A_n}{(1+k)^n}-A_0,$$

$$\sum_{t=1}^{n}\frac{C_t}{(1+k)^t}=\sum_{t=1}^{n}\frac{(1+a_t)A_{t-1}}{(1+k)^t}-(1+k)\sum_{t=0}^{n-1}\frac{A_t}{(1+k)^{t+1}}-\frac{A_n}{(1+k)^n}+A_0.$$

---

[1] A number of issues are sidestepped in this presentation. First, cash flows can be interpreted from different viewpoints, e.g., from the viewpoint of equity interests or all suppliers of capital, but these different meanings of cash flows are not discussed. Second, we do not dwell on the difficulties of defining R_n which can be thought of as "economic value" at the end of the valuation horizon. In finance textbooks, R_n is sometimes referred to as the "horizon" value at the end of period n (Brealey and Myers 1988, 64). Finally, we make the standard assumption that the cost of capital is constant. The analysis would hold for a variable cost of capital but the results would be more complex.

Equivalently:

$$\sum_{t=1}^{n}\frac{C_t}{(1+k)^t}=\sum_{t=1}^{n}\frac{(a_t-k)A_{t-1}}{(1+k)^t}-\frac{A_n}{(1+k)^n}+A_0. \tag{2}$$

Substituting back into equation (1):

$$V_0=A_0+\sum_{t=1}^{n}\frac{(a_t-k)A_{t-1}}{(1+k)^t}+\frac{R_n-A_n}{(1+k)^n}. \tag{3}$$

Equations (2) and (3) are accounting-based valuation models. These equations reflect basic bookkeeping relationships which assume only that the accounting procedures are based on a comprehensive income concept, i.e., clean surplus. Equation (3) shows that the present value of a project is the sum of: (1) initial book value, (2) present discounted value of the project's residual income in period t, $(a_t-k)A_{t-1}$, and (3) the present discounted value of the difference between the terminal economic value of the project and the terminal book value.

Equation (3) is almost identical to *Theorem 1* in Peasnell (1982b, 364) which can be derived by subtracting C_0, the initial cash outflow (when $t=0$ begins the life of the project) or "economic value" (when $t=0$ is an interim period in the life of the project), from both sides of equation (3). Equation (4) in Peasnell (1981, 52) can be derived from equation (3) by assuming that $A_0=C_0$ and $A_n=R_n$. Ohlson (1988, 33), who uses a general equilibrium framework, shows that the equity price of the firm is the sum of its book value plus discounted expected abnormal earnings. Ohlson points out (p. 31) that abnormal earnings can be viewed as residual income; therefore, his derivation can be viewed as a special case of equation (3) when $n=\infty$. Preinreich (1938, 240) derived the same result in continuous time and assumed $A_n=R_n$.

II. Simplification With Constant *ARR* and Constant Growth

Equations (2) and (3) can be simplified if the single-period *ARR*s are constant and a constant-growth rate is assumed. Further simplification is possible for other special cases.

Constant ARR

Several writers who have been concerned with the use of accounting data in the valuation process, e.g., Preinreich (1932, 1937), Fruhan (1979), Wilcox (1984), and Leibowitz and Kogelman (1990), have developed accounting-based DCF models by assuming a constant *ARR*. None of this work discusses the nature of this constant and how to compute it. The general properties of a constant *ARR* are discussed here and the final section of the paper deals with questions relating to its computation.

If a sequence of *ARR*s is constant, the constant must be equal to the value of a that solves:

$$A_0=\frac{C_1}{(1+a)}+\frac{C_2}{(1+a)^2}+\ldots+\frac{C_n}{(1+a)^n}+\frac{A_n}{(1+a)^n}. \tag{4}$$

The proof that a constant accounting rate of return must take the same form as the internal rate of return (*IRR*) is given by Kay (1976, 449).[2] Note that when A_0 equals C_0

---

[2] Throughout this analysis, it is assumed that a is unique. Since the terminal values (R_n) "typically constitute the majority of DCF-generated values" (DeAngelo 1990, 100), the terminal book value (A_n) would, in

(the initial cash outflow or "economic value") and A_n equals R_n (the *horizon* value), a is the same as the *IRR*, r. Unlike some other work related to *ARR*s that is concerned with issues related to profitability analysis, e.g., Edwards et al. (1987, 50–68), the analysis here does not assume that beginning and ending book values also are equal to economic values.

In an unpublished paper, Archer and Peasnell (1982, 2) questioned the significance of a except as a proxy for the *IRR*.[3]

> We cannot see how to assign conceptual significance to an internal rate of return derived from accounting data except as an attempt to operationalise economic IRR. It seems to us that the rate of return concept *per se* is meaningful only in relation to ventures and time periods such that the calculation of the rate of return is based upon a well-determined entry price and final exit price (either of which may be zero), not upon accountants' "pseudo-prices."...

Again, like other researchers in this area, Archer and Peasnell are concerned primarily with the use of the *ARR* to estimate the economic rate of return.

Even though a is based on *pseudo* prices and, therefore, might be described as a *pseudo IRR*, this performance measure has two important properties. First, if a sequence of *ARR*s is constant, the constant must be equal to a, i.e., all $a_t = a$. In other words, whenever an *ARR* is constant, or is assumed to be constant, over any time segment, the *ARR* must be equal to a, as defined in equation (4). This property of a constant *ARR* has not been discussed in the literature that assumes a constant *ARR*.

Second, the quantities, I_t and A_t, are reported in actual or *pro forma* financial statements. In contrast to these reported values, a series of I_t^* and A_t^* can be imputed so that the ratio of I_t^* to A_{t-1}^* is constant and equal to:

$$a = I_t^* / A_{t-1}^*.$$

where I_t^* and A_t^* are similar to the concepts of economic income and economic value, respectively, and these concepts have a long history in accounting.[4] They are most often discussed in connection with the theory of economic income and depreciation where a reference to Hicks (1939) usually can be found. It is shown here that these concepts have a more practical application in financial analysis.

Since any sequence of A_t, I_t, and a_t can be redefined as A_t^*, I_t^*, and a, equation (2) can be rewritten, without loss of generality, as:

$$\sum_{t=1}^{n} \frac{C_t}{(1+k)^t} = (a-k)\sum_{t=1}^{n} \frac{A_{t-1}^*}{(1+k)^t} - \frac{A_n}{(1+k)^n} + A_0. \tag{5}$$

Thus, V_0 becomes:

$$V_0 = A_0 + (a-k)\sum_{t=1}^{n} \frac{A_{t-1}^*}{(1+k)^t} + \frac{R_n - A_n}{(1+k)^n}. \tag{6}$$

Equation (6) is a simplification of equation (3) when the *ARR* is constant.

---

cases where equity valuation was concerned, generally be large enough to insure the uniqueness of a. This conjecture assumes that in cases where cash flows were negative, A_n and R_n would have the same order of magnitude and, therefore, dominate in the caculation of a. For a discussion of issues related to the uniqueness of an internal rate of return, see Teichroew et al. (1965).

[3] The paper is an unpublished comment on Brief et al. (1980).

[4] Note that $A_0 = A_0^*$ and $A_n = A_n^*$.

Further Simplification Based on Constant Growth

Equations (5) and (6) can be simplified by defining g as the growth rate per period of book value from $t=0$ to $t=n$, i.e.,

$$g=\sqrt[n]{\frac{A_n}{A_0}}-1.$$

Therefore:

$$A_n=A_0(1+g)^n$$

is an identity and present discounted cash flows can be rewritten as:

$$\sum_{t=1}^{n}\frac{C_t}{(1+k)^t}=(a-k)\sum_{t=1}^{n}\frac{A^*_{t-1}}{(1+k)^t}+A_0(1-d^n), \tag{7}$$

where:

$$d=\frac{1+g}{1+k}.$$

A further simplification is made possible by assuming that g is also the growth rate of book value in *each* period over the valuation horizon so that:

$$A^*_t=A_0(1+g)^t,$$

for $t=1,2,\ldots,n-1$. In effect, assuming that the A^*_{t-1} grow at a constant rate, g, is equivalent to assuming that income grows at the rate, g, and that cash flows grow at the rate, g. This means the underlying model is in effect, identical to the constant-growth dividend model except that it is expressed in terms of the initial book value, not price.[5]

Substituting $A_0(1+g)^t$ for A^*_t, summing the geometric progression and assuming that $k\neq g$:

$$\sum_{t=1}^{n}\frac{A^*_{t-1}}{(1+k)^t}=\sum_{t=1}^{n}\frac{A_0(1+g)^{t-1}}{(1+k)^t}=A_0\left(\frac{1-d^n}{k-g}\right).$$

Then replacing:

$$\sum_{t=1}^{n}\frac{A^*_{t-1}}{(1+k)^t},$$

by:

$$A_0\left(\frac{1-d^n}{k-g}\right),$$

equation (7) becomes:

$$\sum_{t=1}^{n}\frac{C_t}{(1+k)^t}=A_0\frac{a-k}{k-g}(1-d^n)+A_0(1-d^n). \tag{8}$$

---

[5] Even when the assumption of constant growth is violated, the next section shows that for reasonably "well-behaved" data, the estimate of V_0 based on the constant growth assumption seems to be quite accurate.

Equivalently,

$$\sum_{t=1}^{n} \frac{C_t}{(1+k)^t} = A_0 \frac{a-g}{k-g}(1-d^n), \tag{9}$$

and, adding:

$$\frac{R_n}{(1+k)^n},$$

to equation (9), V_0 becomes:

$$V_0 = A_0 \frac{a-g}{k-g}(1-d^n) + \frac{R_n}{(1+k)^n}. \tag{10}$$

Adding:

$$\frac{R_n}{(1+k)^n},$$

to equation (8), V_0 can be written as:

$$V_0 = A_0 + A_0 \frac{a-k}{k-g}(1-d^n) + \frac{R_n - A_n}{(1+k)^n}. \tag{11}$$

Other Simplifications

The use and limitations of equations (10) and (11) are discussed in the next section. Before dealing with these issues, some other special cases are noted. Some of these special cases have appeared in accounting, economics, and finance literature. One of these was derived almost 60 years ago by Preinreich (1932, 276) in an article in *The Accounting Review.*

$n \to \infty$ and $g < k$. When $n \to \infty$ and $g < k$, it immediately follows from equation (10) that:

$$V_0 = A_0 \frac{a-g}{k-g}. \tag{12}$$

Since $C_1 = A_0(a-g)$, it can be seen that equation (12) is similar to a constant-growth dividend model expressed in terms of accounting quantities.

When V_0 is defined as P_0, the price of a security, and E_1 is earnings in period 1, equation (12) can be redefined in terms of the price/earnings ratio. Since $E_1 = aA_0$:

$$\frac{P_0}{E_1} = \frac{1-g/a}{k-g}. \tag{13}$$

Defining q as the retention rate, g can be written as qa and:

$$\frac{P_0}{E_1} = \frac{1-q}{k-qa}. \tag{14}$$

Equation (14) appeared in Leibowitz and Kogelman (1990, 33).

$R_n/A_n=V_0/A_0$. By implication, $R_n=V_0(1+g)^n$. Substituting $V_0(1+g)^n$ for R_n in equation (10):

$$V_0=A_0\frac{a-g}{k-g}. \tag{15}$$

Equation (15) is the same as equation (12). A similar result also was obtained by Kay (1976, 455–56) who assumed a steady state situation with $k=r$.

$A_n=R_n$. Substituting $A_0(1+g)^n$ for R_n in equation (10) and dividing by A_0:

$$\frac{V_0}{A_0}=\frac{a-g}{k-g}(1-d^n)+d^n. \tag{16}$$

Equation (16) was derived by Preinreich (1932, 276) in a slightly different form. Similarly, substituting $A_0(1+g)^n$ for R_n in equation (11) and dividing by A_0:

$$\frac{V_0}{A_0}=\frac{a-k}{k-g}(1-d^n)+1. \tag{17}$$

Equation (17) appeared in Preinreich (1937, 135).

More recently, the expression:

$$\frac{V_0-A_0}{A_0}=\frac{a-k}{k-g}(1-d^n), \tag{18}$$

appeared in Fruhan (1979) and Wilcox (1984) who were concerned with issues related to the determination of "shareholder value." Equation (18) can be derived from equation (17) by subtracting 1 from both sides of the equation. Fruhan and Wilcox did not consider the more general case when A_n does not equal R_n.

$k=r$. If the cost of capital, k, equals the *IRR*, r, $V_0=C_0$. Then, from equation (11):

$$\frac{C_0-A_0}{A_0}=\frac{(a-r)}{r-g}(1-d^n)+d^n\frac{R_n-A_n}{A_n}. \tag{19}$$

Solving for r:

$$r=a-(r-g)\left(\frac{C_0-A_0}{A_0}\right)+\frac{(r-g)\left(\dfrac{R_n}{A_n}-\dfrac{C_0}{A_0}\right)}{1-d^n}. \tag{20}$$

The second and third terms on the right-hand side of equation (20) can be viewed as the error in using the *ARR* to estimate the internal rate of return. The error depends on the economic/book value ratios at the beginning and at the end of the time horizon. This expression was derived and analyzed in Brief and Lawson (1991a, 1991b).

Other Special Cases. When $a=k$, equation (11) becomes:

$$V_0=A_0+\frac{R_n-A_n}{(1+k)^n}.$$

In this case, the present value equals the initial book value plus an adjustment whose magnitude depends on the extent to which the terminal economic value and accounting book value differ.

When $a=g$, equation (10) can be written as:

$$V_0=\frac{R_n}{(1+k)^n}.$$

When earnings are entirely reinvested, there are no periodic cash flows and the present value of the project depends only on its residual value.

When $k=g$, equations (10) and (11) are not defined. However, $A_0(1+g)^t$ can be substituted for A_t^* in equation (5) which then becomes:

$$\sum_{t=1}^{n}\frac{C_t}{(1+k)^n}=A_0\frac{n(a-k)}{(1+k)}=A_0\frac{n(a-g)}{(1+k)},$$

and

$$V_0=A_0\frac{n(a-g)}{(1+k)}+\frac{R_n}{(1+k)^n}.$$

III. Use and Limitations of Equations (10) and (11) in Financial Analysis

There are two major issues concerning the use of equations (10) and (11) in financial analysis. First, if the single-period *ARR*s are not constant, how can a be computed? Second, suppose the growth of book value is not constant. How much error is introduced by assuming constant growth?

Calculation of a

One might argue that the use of equations (10) and (11) to estimate V_0 is, from a practical viewpoint, not really more direct than using the cash-based DCF model. This is because the calculation of a from equation (4) is based on cash flows. Therefore, if cash flows must be known to calculate a, using equation (1) to estimate V_0 would actually be more direct than using equations (10) and (11).[6] Thus, if equation (4) has to be used to determine a, the argument that the accounting-based valuation method is circular would have merit. However, a can be estimated directly from the sequence of single-period *ARR*s.

Insight into the nature of estimates of a from a sequence of *ARR*s can be obtained from the relationship between a and a weighted average of single-period *ARR*s. This relationship was first derived by Kay (1976, 452–53) and then elaborated on by Edwards et al. (1987, 37). Assuming the comprehensive income concept is employed, a can be defined for any time segment in the life of a project as the weighted average of actual single-period rates of return, a_t.[7] The definition of the constant, a, in equation (4) is the algebraic equivalent of the value of a that solves:

$$a=\sum_{t=1}^{n}w_t a_t, \tag{21}$$

---

[6] If the comprehensive income concept is employed, cash flows can, of course, be determined directly from accounting income, I_t, and book value, A_t, since by definition, $C_t=I_t-(A_t-A_{t-1})$.

[7] In their discussion, Edwards et al. were concerned with using a to estimate the *IRR* in both multi-period and single-period situations. Consequently, much of their discussion was devoted to issues related to how A_0 and A_n should be revalued so that the resulting rate of return had economic significance.

where the weights, w_t, are:

$$w_t=\frac{\dfrac{A_{t-1}}{(1+a)^t}}{\displaystyle\sum_{j=1}^{n}\frac{A_{j-1}}{(1+a)^j}},$$

and

$$\sum_{t=1}^{n} w_t=1.$$

Thus, given A_t and a_t, a can be determined by solving equation (21) for a; but using this method also is equivalent to using equation (4) to calculate a. Therefore, once again, there seems to be circularity in the argument. However, the definition of a as a weighted average *ARR* is important because it provides insight into the nature of the bias in using more direct ways to transform a sequence of *ARR*s into an estimate of a.

Peasnell (1982a, 29) comments that "knowledge of the general weight function ... is often informative" since, normally, there are restrictions on the w_t.[8] Under most circumstances, the book values, A_t, are positive. Therefore, since the weights sum to 1 and must be between 0 and 1, the analyst knows that in any sequence of accounting returns, the minimum and maximum values of a_t define the range of values that a might take.

Furthermore, when the *ARR*s are random with no time trend, a simple average of the *ARR*s will in expectation be equal to a. If the time trend of *ARR*s is positive, the simple average will overstate a if $g>a$ and understate a if $g<a$. If the time trend is negative, the opposite relationships will hold.[9]

The examples in table 1 show the relationship between a simple arithmetic average of the single-period *ARR*s and the *pseudo IRR* for different assumptions about the time series of a_t. All of the cases assume the same cash flows and opening and closing book values. Therefore, the *pseudo IRR* is the same in all cases and $a=0.12$. In the first case, which is taken directly from the example in Brealey and Myers (1988, 63), the *ARR* in each period equals 0.12. Case 2 illustrates "random" *ARR*s, varying from 0.09 to 0.15. The mean is 0.121 which is very close to a. In the next two cases the *ARR*s are upward and downward sloping time series with means of 0.117 and 0.125. Again, the means are close to a. Only in the last case, which displays a very unstable sequence of *ARR*s, is there a big difference between the arithmetic mean of the sequence of the a_t and the *pseudo IRR*.

These examples suggest that when the time series of a_t is reasonably stable, the bias in estimates of a based on a simple arithmetic mean of *ARR*s will not be large.[10] In practice, when dealing with the valuation of large units like a division of a firm or the

---

[8] Peasnell's concern here, as indicated earlier, was to estimate the *IRR*, not a, although much of his analysis is applicable to the questions being considered here.

[9] In Kay's (1976, 454) discussion of the relationship between the simple average and a over the life of a firm, he implicitly assumed that $g<a$ which is, of course, quite reasonable; however, over a limited time segment, g might be greater or less than a.

[10] Obviously, the bias in estimates of a using a simple average needs to be studied in more detail. The examples here are only suggestive of the amount of bias that might be expected.

Table 1
Effect of *ARR* Variability on Average *ARR*
($a=0.12$)
(millions of dollars)

Year	0	1	2	3	4	5	6
Cash Flows		−0.80	−0.96	−1.15	−1.39	−0.20	−0.23
Book Value	10.00	—	—	—	—	—	26.47
Case 1: Constant ARRs			Average *ARR*=0.120				
ARR		0.120	0.120	0.120	0.120	0.120	0.120
Income		1.20	1.44	1.73	2.07	2.49	2.89
Book Value	10.00	12.00	14.40	17.28	20.74	23.43	26.47
Case 2: "Random" ARRs			Average *ARR*=0.121				
ARR		0.140	0.100	0.150	0.110	0.090	0.135
Income		1.40	1.22	2.16	1.95	1.89	3.12
Book Value	10.00	12.20	14.38	17.69	21.02	23.11	26.47
Case 3: Increasing ARRs			Average *ARR*=0.117				
ARR		0.070	0.080	0.100	0.130	0.140	0.181
Income		0.70	0.92	1.34	2.06	2.70	4.01
Book Value	10.00	11.50	13.38	15.87	19.32	22.23	26.47
Case 4: Decreasing ARRs			Average *ARR*=0.125				
ARR		0.200	0.165	0.130	0.100	0.085	0.072
Income		2.00	2.11	2.06	1.91	1.90	.175
Book Value	10.00	12.80	15.87	19.09	22.38	24.49	26.47
Case 5: Unstable ARRs			Average *ARR*=0.196				
ARR		1.174	0	0	0	0	0
Income		11.74	0	0	0	0	0
Book Value	10.00	22.54	23.50	24.65	26.04	26.24	26.47

The data on cash flows and opening and closing book values, and the data in case 1, are taken from Brealey and Myers (1988, 63).

firm itself over a short time horizon, the *ARR* would be expected to be reasonably stable. Indeed, if the *ARR*s displayed a great deal of instability, it would be doubtful if the analyst would, either directly or indirectly, base DCF analysis on such data.

Error in Assuming Constant Growth of Book Value

Another estimation error is introduced by assuming that book value grows at a constant rate. The error in assuming that $A_t^* = A_0(1+g)^t$ can be evaluated as follows. Define:

$$S = \sum_{t=1}^{n} \frac{A_{t-1}^*}{(1+k)^t},$$

and

$$s=A_0\frac{(1-d^n)}{(k-g)},$$

where S is the exact value and s is the approximate value of:

$$\sum_{t=1}^{n}\frac{A^*_{t-1}}{(1+k)^t}.$$

Further, let:

$$x=\frac{s}{S}.$$

Thus, equation (5), which is the *exact* value of:

$$\sum_{t=1}^{n}\frac{C_t}{(1+k)^t},$$

can be written as:

$$EXACT=S(a-k)+xS(k-g),$$

and equation (8), which is the *approximate* value of:

$$\sum_{t=1}^{n}\frac{C_t}{(1+k)^t},$$

can be written as:

$$APPROXIMATE=xS(a-k)+xS(k-g).$$

Therefore,

$$\frac{APPROXIMATE}{EXACT}=\frac{x(a-k)+x(k-g)}{(a-k)+x(k-g)}.$$

If we now assume that $a\neq k$ and $k\neq g$ and let:

$$k-g=y(a-k),$$

then the accuracy of the approximation is:

$$\frac{APPROXIMATE}{EXACT}=\frac{x+xy}{1+xy}. \qquad (22)$$

In preliminary work, it was apparent that even if the A^*_t display substantial variability about a trend line of constant growth, the accuracy of the approximation will, nevertheless, be quite high. For example, if $n=10$, $A_0=10$, and $A_n=20$, the growth rate, g, will be equal to 0.0718. Let:

$$A^*_t=10(1.0718)^t+e_t,$$

where $e_t=-8$ in periods $t=1$, 2, 5, 6, and 9 and $+8$ in periods $t=3$, 4, 7, and 8. If

Table 2
Accuracy of Approximation of Present Value of Cash Flows
(x=1.0911)

xy	$\frac{x+xy}{1+xy}$	xy	$\frac{x+xy}{1+xy}$
5	1.02	−5	.98
4	1.02	−4	.97
3	1.02	−3	.95
2	1.03	−2	.91
1.5	1.04	−1.5	.82
1	1.05	−1	Not Defined
.5	1.06	−.5	1.18
.2	1.08	−.2	1.11
.1	1.08	−.1	1.10
.05	1.09	−.05	1.10
.01	1.09	−.01	1.09

$\frac{x+xy}{1+xy}$ = ratio of approximation of the present discounted value of cash flows to the exact value of the present discounted values of cash flows.

$k=0.10$, then $S=74.3275$, and $s=81.0987$, and

$$x=\frac{81.0987}{74.3275}=1.0911.$$

Table 2 shows the accuracy of the approximation of the present discounted value of cash flows when $x=1.0911$, calculated from equation (22) for various values of xy. The error in the approximation is under 10 percent except for negative values of xy in the region between about $xy=-2$ and $xy=-0.1$. It should be borne in mind that since x is positive, xy is negative only when y is negative and this will occur when $g>k$ or $k>a$. Neither of these conditions is likely to occur with great frequency. Therefore, in most "normal" situations, y and xy are positive and x can be viewed as a lower bound on the accuracy of the approximation.

Example

The data on cash flows and opening and closing book values, and the data in case 1, in table 1 are taken from Brealey and Myers (1988, 63). The accounting rate of return in this example is constant and, therefore, equal to the *pseudo IRR*. Brealey and Myers calculate present value by computing the present value of free cash flows for years 1 to 6 and then add to this amount the present discounted value of the horizon value which is the present value of the stream of cash flows after year 6. In the example, free cash flow in period 7 equals 1.59 and is assumed to grow indefinitely at 6 percent per period. Also, the cost of capital is assumed to be $k=0.10$. The net present value of 18.8 million dollars is obtained using equation (1) by calculating:

$$\sum_{t=1}^{n}\frac{C_t}{(1+k)^t}=\frac{-0.80}{1.1}+\frac{-0.96}{(1.1)^2}+\frac{-1.15}{(1.1)^3}+\frac{-1.39}{(1.1)^4}+\frac{-0.20}{(1.1)^5}+\frac{-0.23}{(1.1)^6}=-3.6,$$

and

$$\frac{R_n}{(1+k)^n}=\frac{1}{(1.1)^6}\frac{1.59}{0.10-0.06}=22.4.$$

Therefore,

$$V_0=-3.6+22.4=18.8.$$

An approximation of the present discounted value of cash flows from periods 1 to 6 can be obtained directly from the accounting data by substituting $k=0.10$, $a=0.12$, $A_0=10$, and

$$g=\sqrt[6]{\frac{A_n}{A_0}}-1=\sqrt[6]{\frac{26.47}{10}}-1=0.176$$

into equation (10). Thus,

$$\sum_{t=1}^{n}\frac{C_t}{(1+k)^t}=10\frac{(0.12-0.176)}{(0.10-0.176)}\left[1-\left(\frac{1.176}{1.10}\right)^6\right],$$

which gives:

$$\sum_{t=1}^{n}\frac{C_t}{(1+k)^t}=-3.6.$$

Since the example assumes that after period 6, $g=0.06$, and an infinite time horizon is assumed, terminal value can be calculated in the same way that the present discounted value of cash flows is calculated, i.e.,

$$\frac{R_n}{(1+k)^n}=\frac{26.47}{(1.10)^6}\frac{(0.12-0.06)}{(0.10-0.06)}=\frac{39.71}{(1.10)^6}=22.4.$$

The accounting-based method to calculate V_0 yields the same results as the method based on cash flows.

IV. Conclusion

Research on *ARR*s has more than a 25-year history and most of it has focused on the limitations of using the *ARR* to make inferences about a firm's profitability. In this paper, the focus is on valuation, not profitability, and a general accounting-based DCF formula that assumes only that accounting policies are based on the comprehensive income concept is derived. The extant accounting system is taken as given and the main concern of this paper is on how to value a project over a limited time slice using accounting data.

Without loss of generality, the model is simplified by replacing the variable, single-period *ARR*s with a constant. Then a constant-growth assumption is made and several other special cases are also discussed.

The results can be viewed as basic relationships that are derived from a double-entry system. Students of accounting should be more familiar with these analytical properties. On a more practical side, financial analysts can base DCF valuations on accounting data instead of following the current practice of using historical accounting data to forecast future earnings and then to estimate cash flows from earnings. The analysis does not resolve all of the relevant issues—clearly, more work needs to be done in assessing the biases inherent in the estimation procedures suggested—but it does suggest that the *ARR* has a new role to play.

The *pseudo IRR* has a straightforward interpretation and useful properties. In previous work, the *pseudo IRR* has been discussed as a proxy for the *IRR*. In this paper, it is discussed as a parameter that enters into the valuation process. This parameter depends only on opening and closing book values and cash flows, not on accounting numbers in interim periods.

This has implications for conventional accounting theory which is, in effect, based on a one-period model. Conventional theory does not address questions about multi-period analysis where the determination of assets and liabilities at the beginning and at the end of the multi-period time horizon appears to be the overriding accounting issue.[11] Thus, when two or more periods are analyzed, only the cash flow series and the change in book value (which sum up to total income over the time horizon) are relevant, not the separate amounts of income reported in the individual periods. If cash flows and the change of book value are specified over any time segment, there are infinitely many *ARR* sequences consistent with that specification whereas there is only one value of the *pseudo IRR*.[12]

While accountants and financial analysts often talk about how accounting data are used in the decision making process, such discussions are characterized by generalities rather than specifics. Even though the DCF techniques used to value equity investments are often accounting-based valuation approaches, the link between accounting numbers and the calculation of present value is, in practice, indirect. Understanding how DCF analysis can be based directly on accounting data leads to a greater appreciation of the general nature of accounting and also provides a compelling reason to give the *pseudo IRR* a more prominent place in financial statement analysis.

[11] From this viewpoint, revenue and expense determination in the interim periods are relevant only insofar as revenues and expenses affect beginning and ending book values. Edwards et al. (1987, 32) make a similar point when they comment that "... *focusing attention on the opening and closing valuations of capital employed* ... serves to emphasize the issue that lies at the heart of the debate on the relevance of accounting profitability."

[12] See footnote 2.

References

Archer, S., and K. V. Peasnell. 1982. Cumulative financial statements. Unpublished paper. University of Lancaster. United Kingdom.

Brealey, R. A., and S. C. Myers. 1988. *Principles of Corporate Finance*. 3d ed. New York: McGraw-Hill.

Brief, R. P., and R. A. Lawson. 1991a. Approximate error in using accounting rates of return to estimate economic returns. *Journal of Business Finance & Accounting* 18 (January): 13–20.

——, and ——. 1991b. Approximate error in using accounting rates of return to estimate economic returns: A correction. *Journal of Business Finance & Accounting* 18 (November): 915–16.

——, B. D. Merino, and I. Weiss. 1980. Cumulative financial statements. *The Accounting Review* 55 (July): 480–90.

DeAngelo, L. E. 1990. Equity valuation and corporate control. *The Accounting Review* 65 (January): 93–112.

Edwards, J., J. Kay, and C. Mayer. 1987. *The Economic Analysis of Accounting Profitability*. Oxford, United Kingdom: Clarendon Press.

Fisher, F., and J. J. McGowan. 1983. On the misuse of accounting rates of return to infer monopoly profits. *American Economic Review* 73 (March): 82–97.

Foster, G. 1986. *Financial Statement Analysis*. 2d ed. Englewood Cliffs, NJ: Prentice-Hall.

Fruhan, W. E., Jr. 1979. *Financial Strategy*. Homewood, IL: Richard D. Irwin.

Harcourt, G. C. 1965. The accountant in a golden age. *Oxford Economic Papers* 5 (March): 66–80.

Hicks, J. R. 1939. *Value and Capital*. Oxford, United Kingdom: Clarendon Press.

Kay, J. A. 1976. Accountants, too, could be happy in the golden age: The accountants rate of profit and the internal rate of return. *Oxford Economic Papers* 17 (November): 66–80.

Leibowitz, M. L., and S. Kogelman. 1990. Inside the P/E ratio: The franchise factor. *Financial Analysts Journal* 46 (November/December): 17–35.

Ohlson, J. A. 1988. Accounting earnings, book value, and dividends: The theory of the clean surplus equation (part I). Unpublished manuscript. New York: Columbia University.

Peasnell, K. V. 1981. On capital budgeting and income measurement. *Abacus* 17 (June): 52–67.

———. 1982a. Estimating the internal rate of return from accounting profit rates. *The Investment Analyst* (April): 26–31.

———. 1982b. Some formal connections between economic values and yields and accounting numbers. *Journal of Business Finance & Accounting* 9 (Autumn): 361–81.

Preinreich, G. A. D. 1932. Stock yields, stock dividends and inflation. *The Accounting Review* 7 (December): 271–89.

———. 1937. Fair value and yield of common stock. *The Accounting Review* 12 (June): 130–40.

———. 1938. Annual survey of economic theory: The theory of depreciation. *Econometrica* 6 (July): 219–41.

Rappaport, A. 1986. *Creating Shareholder Value: The New Standard for Business Performance*. New York: The Free Press.

Salamon, G. L. 1985. Accounting rates of return. *American Economic Review* 75 (June): 495–504.

Solomon, E. 1966. Return on investment: The relationship of book-yield to true yield. In *Research in Accounting Measurement*, edited by R. K. Jaedicke, Y. Ijiri, and O. Nielsen. American Accounting Association.

Teichroew, D. A., A. A. Robichek, and M. Montalbano. 1965. An analysis of criteria for investing and financing decisions under certainty. *Management Science* 12 (November): 151–80.

Whittington, G. 1988. The usefulness of accounting data in measuring the economic performance of firms. *Journal of Accounting and Public Policy* 7 (Winter): 261–66.

Wilcox, J. W. 1984. The P/B-ROE valuation model. *Financial Analysts Journal* 40 (January–February): 58–66.

THE ACCOUNTING REVIEW
Vol. 67, No. 2
April 1992
pp. 427–439

The Effect of *Ex Ante* Earnings Uncertainty on Earnings Response Coefficients

Eugene A. Imhoff, Jr.
The University of Michigan
Gerald J. Lobo
Syracuse University

SYNOPSIS AND INTRODUCTION: This study examines the effect of the uncertainty in analysts' earnings forecasts on the relation between unexpected returns and unexpected earnings. Numerous theorists have considered the effect of uncertainty on firm value, with particular interest in the uncertainty in a firm's future cash flows that underpin firm value. Since accrual accounting earnings represent a theoretical proxy for future cash flows, the effect of earnings uncertainty on firm value is also of considerable interest. However, observed uncertainty in accounting earnings may be attributable to noise (garbling) in the earnings signal or to the fundamental uncertainty of expected future cash flows, or both. Moreover, theory suggests these different forms of uncertainty may have differing effects on firm value. To date, there is little empirical evidence concerning the effect of uncertainty in earnings on firm value.

We investigate the effect of *ex ante* earnings uncertainty by using the familiar linear relation between unexpected stock returns (*UR*) and unexpected earnings (*UE*), with 3,167 firm-year observations collected over the six-year period 1979–84. The variance in analysts' earnings forecasts just prior to a firm's annual earnings announcement is employed as our firm-specific proxy for *ex ante* uncertainty. Our results indicate a systematic relation between *ex ante* uncertainty and the information content of earnings. A given unit of earnings news has a greater effect on unexpected stock price change as the amount of pre-earnings-announcement

The authors wish to thank the participants in workshops at The University of Michigan, Purdue University, University of Iowa, University of Connecticut, University of Illinois, and University of Notre Dame for their helpful comments on earlier drafts of this paper. We are also grateful for the comments from and discussions with Jeff Abarbanell, Vic Bernard, Tom Carroll, Art Goldberger, Florence Kirk, Bill Lanen, Tom Linsmeier, Robert Lipe, Jim Noel, Gordon Potter, and the referees, and the research assistance of Chai Kim and Pete vanAmson. Financial support was provided by the School of Business, The University of Michigan.

Submitted September 1989.
Accepted February 1991.

uncertainty decreases. Firms with relatively high *ex ante* uncertainty exhibit little or no systematic price change at the time earnings are announced.

Sensitivity tests reveal our results to be robust over numerous alternative specifications of the variables and models employed. Sensitivity tests also suggest that our results are not driven by either firm size or the amount of information available about the firm. In addition, we develop and report results of a model that controls for the effects of uncertainty. These results imply that the dispersion (disagreement) in analysts' earnings forecasts is more likely to be a proxy for noise in the financial reporting system than a proxy for fundamental uncertainty in a firm's future cash flows.

Key Words: *Analysts' forecasts, Earnings uncertainty, Earnings response coefficients, Market reaction.*

Data Availability: *Data used in this study are available from the authors and from the Institutional Brokers Estimate System (I/B/E/S) compiled by Lynch, Jones & Ryan of New York.*

THIS first section briefly discusses the related research to date. It is followed by a description of the measures and research methodology. Section III provides the results and an explanation of the sensitivity tests conducted and is followed by a model that controls for *ex ante* uncertainty. The final section discusses the implications and presents the conclusions of the study.

I. Related Research

Current accounting research includes numerous studies that attempt to identify firm-specific characteristics that are likely to differentially affect the *UE-UR* relation. These factors include: the differential quality of preannouncement information (Verrecchia 1980), firm size (Atiase 1985; Freeman 1987), the exchange on which a firm's stocks are traded (Grant 1980; Atiase 1987), and the effect of earnings predictability (Lipe 1990).

Theory also suggests that uncertainty should systematically affect changes in firm value. Verrecchia (1980), Holthausen and Verrecchia (1988), Robichek and Myers (1966), and Epstein and Turnbull (1980) have all considered the theoretical impact of uncertainty on firm value. These studies suggest that greater fundamental uncertainty about a firm's future cash flows will cause larger stock price responses to value relevant information, *ceteris paribus*. Although nonearnings information may be value-relevant, the relation between earnings and firm value views earnings as a source of information for predicting future cash flows. Hence, greater uncertainty in earnings may indicate greater uncertainty in future cash flows, thereby resulting in larger stock price responses per unit of (value-relevant) earnings news.

Alternatively, earnings uncertainty may represent noise in the process generating the earnings signal. The subjective nature of the accrual accounting model permits the firm's managers to influence reported earnings without necessarily altering its fundamental economic flows, currently or in the future. Holthausen and Verrecchia (1988) define "noise" as that portion of the signal with no information content. Hence, if earnings uncertainty is a manifestation of noise, greater uncertainty in earnings should result in smaller stock price responses per unit of earnings news. In the limit, the noise could be so great that earnings is no longer value-relevant, at least in its reported form.

Since earnings uncertainty may have a different effect on earnings response coefficients (ERCs), depending on whether it reflects fundamental uncertainty or noise or both, it would be desirable to identify a pure proxy for each form of uncertainty. Because no pure proxies are available, all variables representing earnings uncertainty are potentially mixtures of fundamental uncertainty and noise. In addition, empirical work has no means of controlling for the relative uncertainty in nonearnings information, which has been acknowledged as being potentially valuation relevant (see, e.g., Holthausen and Verrecchia 1988; Lipe 1990). The problems created by these control issues may help to explain the inconclusive results of some related empirical work on the effects of uncertainty (see, e.g., Lang 1989; Rao 1989; Pincus 1983).

The measure of *ex ante* earnings uncertainty employed here is the dispersion in analysts' forecasts, which may also be viewed as a measure of earnings predictability. Both Pincus (1983) and Lipe (1990) have examined the impact of earnings predictability (uncertainty) on firm value.[1] Lipe (1990) provides evidence that greater predictability (less uncertainty) in earnings results in larger ERCs. To the extent that our proxy for *ex ante* earnings uncertainty is also a measure of earnings predictability, its impact on ERCs could be expected to behave like noise in earnings rather than fundamental uncertainty.[2]

Because of the relation between *ex ante* earnings uncertainty and earnings predictability, this study may be viewed as a complement to Lipe's study. Lipe (1990) uses the standard deviation of residuals from a first-order autoregressive process of past earnings as a measure of predictability (uncertainty) in a time-series association test framework. This study uses the mean and the standard deviation of analysts' forecasts, which may incorporate contemporaneous information about a firm's expected earnings, in a cross-sectional event-study framework.

Our research hypothesis is that *ex ante* earnings uncertainty has a systematic effect on the market's response to earnings news. Moreover, on the basis of evidence provided by Kormendi and Lipe (1987) and Lipe (1990), greater *ex ante* uncertainty is expected to result in smaller earnings response coefficients. At the same time, we recognize that the inability to control for the potential effects of fundamental uncertainty and uncertainty in nonearnings information are limiting factors that could confound tests of the research hypothesis. To the extent that there may be offsetting effects due to fundamental changes in expected future cash flows reflected in our earnings-uncertainty proxy, it is less likely that systematic effects will be observed in our tests.

---

[1] While pure empirical tests of theory regarding the effects of uncertainty (predictability) have not been possible, the results reported by Pincus (1983) were inconsistent with the theory.

[2] Research results not yet reported (Imhoff and Lobo 1991) document a very strong relation between I/B/E/S forecast errors and the dispersion of analysts' forecasts, with greater uncertainty (in all fiscal quarters) resulting in larger forecast errors, across firms and over time.

II. Measurements and Methodology

The empirical tests of the research hypothesis require measures of three variables: unexpected stock returns (UR), unexpected annual earnings (UE), and dispersion of analysts' earnings forecasts.

Measuring Unexpected Returns

Unexpected stock returns are derived from the market model:

$$R_{i\tau}=a_i+b_iR_{m\tau}+\epsilon_{i\tau}, \tag{1}$$

where:

$R_{i\tau}$ = return on firm i on day τ,
$R_{m\tau}$ = return on the CRSP value-weighted market index on day τ,
$\epsilon_{i\tau}$ = a stochastic disturbance term, and
a_i and b_i = intercept and slope coefficient, respectively, for firm i.

The coefficients, a_i and b_i, are ordinary least squares estimates using daily returns from day -202 to day -3.[3] These estimates, $\hat{a}_i$ and $\hat{b}_i$, are then used to compute UR_{it}, the cumulative unexpected security returns during the earnings announcement period, as follows:

$$UR_{it}=\sum_{t=-1}^{0}[R_{it}-(\hat{a}_i+\hat{b}_iR_{mt})]. \tag{2}$$

The results reported are for cumulative two-day returns on the day of and the day preceding the earnings announcement in *The Wall Street Journal*.

Measuring Unexpected Earnings

Unexpected earnings are measured as follows:

$$UE_{is}=\frac{A_{is}-F_{is}}{P_{is}}, \tag{3}$$

where:

A_{is} = actual earnings for firm i in period s,
F_{is} = expected annual earnings for firm i in period s, measured as the mean security analysts' annual earnings forecast on the I/B/E/S tape in the month immediately preceding that of the earnings announcement, and
P_{is} = stock price for firm i two trading days before the annual earnings announcement date.

The Institutional Brokers Estimate System (I/B/E/S) database, compiled by Lynch, Jones & Ryan, contains monthly summary statistics, such as means, medians, and standard deviations of earnings per share (EPS) forecasts, contributed by more than 50 brokerage firms for more than 2,000 publicly owned companies. The number of analysts providing forecasts for a given firm and the number of firms for which forecasts are available vary by month and year.[4]

---

[3] The minimum number of nonmissing estimation period returns for any firm-year observation was 197.

[4] For a further discussion of the nature of I/B/E/S forecasts and their potential limitations, see Cornell and Landsman (1989).

Different measures of unexpected earnings may alter the interpretation of tests regarding the information content of earnings announcements. The mean of analysts' forecasts of EPS (as reported by I/B/E/S) in the month preceding the earnings announcement is used as our proxy for expected earnings. For the reported results, the forecast error $(A-F)$ is deflated by the stock price two days before the annual earnings announcement in *The Wall Street Journal.* The stock price deflator is suggested by Christie (1987) as a way to reduce scale effects. This measure expresses unexpected earnings per share in return form as the unexpected earnings per dollar of common equity.[5] With UE as the independent variable, the following equation is estimated:

$$UR_{it}=\alpha+\beta UE_{it}+\epsilon_{it}. \tag{4}$$

This model is similar to those reported in numerous prior studies and is employed here to enhance comparability with prior work.

Data Requirements and Sample Description

The most restrictive data requirement was the availability of a measure of uncertainty in earnings forecasts. The I/B/E/S database was used to obtain measures of *ex ante* uncertainty, which is designated hereafter as ASD and which equals the firm-specific standard deviation of the forecasts for each period deflated by stock price two days prior to the earnings announcement. Our initial sample consisted of all COMPUSTAT companies with annual earnings announcement dates and with five or more analysts' earnings forecasts reported on I/B/E/S in the month prior to their actual annual earnings announcement for the six-year test period 1979–84. This restriction was imposed to permit a reasonable statistical estimate of the standard deviation of analysts' forecasts for ASD. The number of forecasts per firm ranged from five to 38, with a median (mean) number of analysts' forecasts per firm of 12 (12.94).

In addition to I/B/E/S data, the sample firms were required to have CRSP daily returns available for the period (−202 to +1), where day zero represents the annual earnings announcement date. The effects of the various data requirements on sample size are summarized in table 1. The final sample consisted of 3,167 firm-year observations for 697 firms during the period 1979–84. Since most companies use a calendar year, 2,279 firm-year observations (72 percent) were from the months of January and February. The cross-sectional tests are based on many different two-day return windows for any given year in our time series, which reduces the effects of cross-sectional dependence in the abnormal returns.[6]

Model and Tests

We first examine the basic equation (4) relation between unexpected earnings and unexpected returns for comparison with earlier work. Next, the sample is divided into three strata based on ranked values of ASD. Firms are ranked each year in descending

---

[5] We implicitly assume no difference in the market value of a dollar of unexpected earnings per share across firms. Evidence provided by Kormendi and Lipe (1987) and Easton and Zmijewski (1989) suggests this is incorrect. Moreover, it is intuitively unappealing since it implies constant cross-sectional P/E ratios. Since P/E ratios vary cross-sectionally, this deflator may be contributing to the variance of the response coefficient (slope) in equation (4). Despite these objections, we employ the price deflator in most of our reported results because the economic interpretation of the response coefficients is enhanced when price per share is used to deflate unexpected earnings per share.

[6] In addition, cross-sectional correlation effects should be minimal in this study because: (1) there is a wide range of industries in the sample, and (2) cross-sectional correlation effects are not severe over short holding periods, two days in this study (Bernard 1987).

Table 1
Sample Selection

Data Requirement	*Total*	*1979*	*1980*	*1981*	*1982*	*1983*	*1984*
	Number of Firm-year Observations Available						
1. Firms with annual earnings announcement dates on COMPUSTAT and five or more I/B/E/S earnings forecasts in the month prior to the announcement month.	3,820	527	590	667	684	721	631
2. Firms from (1) above on CRSP daily price tape.	3,594	510	570	637	652	683	542
3. Firms from (2) above with complete price, return, price adjustment, and EPS data.	3,167	437	492	545	561	593	539

Summary

(1) Total I/B/E/S firms.	5,485
(2) Firms in (1) with ≥5 forecasts in any month.	2,183
(3) Firms in (2) that are on the COMPUSTAT quarterly and annual tapes.	1,132
(4) Firms in (3) that are on the CRSP daily tape.	1,047
(5) Firms in (4) with data available on all tapes (earnings announcement date, ≥5 forecasts in month prior to announcement, annual EPS, daily returns, prices).	697

order of ASD. Each ranked array is then divided into three equal-sized strata: high, medium, and low uncertainty. We test the research hypothesis by examining the effect of *ex ante* uncertainty on the earnings response coefficients for the two-day period (−1, 0) surrounding the annual earnings announcement date by the following model:

$$UR_{it}=\alpha_1+\alpha_2 D_M+\alpha_3 D_H+\beta_1 UE_{it}+\beta_2 D_M UE_{it}+\beta_3 D_H UE_{it}+\epsilon_{it}, \qquad (5)$$

where:

D_M = a dummy variable equal to 1 when a firm-year observation is in the MIDDLE uncertainty stratum for that year, and 0 otherwise;
D_H = a dummy variable equal to 1 when a firm-year observation is in the HIGH uncertainty stratum for that year, and 0 otherwise;
and all other variables are defined earlier.

This model is formulated to permit the intercept and response coefficient to vary across strata. α_1 and β_1 represent the intercept and response coefficient, respectively, for the low ASD stratum. For the middle and high strata, the intercepts are $\alpha_1+\alpha_2$ and $\alpha_1+\alpha_3$, and the response coefficients are $\beta_1+\beta_2$ and $\beta_1+\beta_3$, respectively. Hence, α_2 and β_2 (α_3 and β_3) represent the differences in the intercepts and response coefficients, respectively, between the middle (high) and low uncertainty strata. Consistent with the research hypothesis, we test the following:

H_{0_1}: $\beta_2=\beta_3=0$, H_{A_1}: $\beta_2<0$; $\beta_3<0$, and
H_{0_2}: $\beta_2-\beta_3=0$, H_{A_2}: $\beta_3<\beta_2$.

Table 2
Relation between Unexpected Earnings (*UE*) and Unexpected Returns (*UR*) for the Period 1979–84

Period	Number of Portfolios	$UR_{is}=\alpha+\beta UE_{is}+\epsilon_{is}$[a] α	 β	 Adjusted R^2	White's χ^2 (Significance)
1979–84	443	−.00084 (−1.343)[b]	.04694 (3.961)[b]	.0322	2.4164 (.2987)

[a] UR_{is} = unexpected two-day return of portfolio i for the day before and the day of the annual earnings announcement (−1, 0);

UE_{is} = unexpected earnings for portfolio i, measured as the difference between actual earnings and the mean I/B/E/S forecast deflated by stock price two trading days before the earnings announcement;

α,β = intercept and slope (response) coefficient; and

ϵ_{is} = residual for portfolio i in period s.

[b] The test statistics in this table are t-statistics because the White's test for heteroskedasticity in the residuals (reported as a χ^2-statistic in the last column) revealed no significant heteroskedasticity. We provide the White's χ^2-statistic in all other tables where regression results are reported. When there is (is not) significant heteroskedasticity, the test statistics reported for the intercepts and slopes are the z-statistics (t-statistics), appropriate for such cases.

On the basis of the prediction of the research hypothesis, we expect differences in response coefficients between the middle (high) and low uncertainty strata to be negative (i.e., $\beta_2<0$ and $\beta_3<0$). Furthermore, we expect the response coefficient for the high ASD stratum to be less (more negative) than the response coefficient for the middle ASD stratum (i.e., $\beta_2>\beta_3$). In other words, if the impact of *ex ante* uncertainty is systematic, the effects (i.e., lower response coefficients) should be greater as uncertainty increases.

III. Results

We estimate equations (4) and (5) in order to assess the effect of uncertainty on the relation between *UR* and *UE*. The test results are based on portfolio data constructed by ranking all firms each year within each of the three uncertainty strata on the basis of their magnitude of *UE*.[7] These three ranked arrays of firms formed each year were then divided into portfolios with a minimum of seven firm-year observations in each.[8] The average values of *UR* and *UE* were then computed for each portfolio to represent the portfolio *UR* and *UE* and to estimate the parameters in equations (4) and (5).

---

[7] The typical explanatory power of models (4) and (5) with pooled time-series and cross-sectional firm-specific data for two-day return periods is low ($R^2 \leq 0.02$ usually). Grouping firms into portfolios can result in a considerable increase in the explanatory power of the regressions (Kmenta 1971, 328). Although grouping observations tends to increase R^2, estimation with grouped observations is less efficient than that based on individual observations because information about variation of the observations within each group is lost. This loss in efficiency will be small if the variation within groups is small relative to the variation across groups. An advantage of grouping is that measurement error will be reduced if the errors associated with individual observations in each group are less than perfectly correlated. We repeated all of the tests reported here on a firm-by-firm basis, with and without outliers, and obtained results completely consistent with those reported here. These additional results are available upon request.

[8] Some portfolios were comprised of eight firm-year observations rather than seven. This difference in portfolio sizes is one source of heteroskedasticity. We report tests of heteroskedasticity along with our other results.

Regression Results

After forming 443 portfolios from the 3,167 firm-year observations (as described above), we estimated equation (4) using ordinary least squares (OLS) regression. Table 2 reports a summary of the OLS regression results for equation (4) for the two-day period surrounding the earnings announcement date. The earnings response coefficient (β) for equation (4) during the two-day announcement period $(-1, 0)$ is positive and significant with a *t*-statistic of 3.961. Because of concern for heteroskedasticity in the residuals, we tested the null hypothesis that the errors are homoskedastic, using the procedure developed by White (1980). The χ^2-statistics and their significance levels are reported for all regressions. Whenever the null was rejected, we employed White's general heteroskedasticity corrected covariance matrix estimates and reported the resulting *z*- and χ^2-statistics in place of *t*- and *F*-statistics.

Next, we consider the differences in response coefficients across the three strata by estimating equation (5). The pooled time-series results are reported in table 3 along with year-by-year results for portfolios ranked on the basis of ASD.

The estimated response coefficients for the three uncertainty strata reported in the pooled results are:

	Response Coefficients
Low uncertainty (β_1)	0.77
Middle uncertainty ($\beta_1+\beta_2$)	0.42
High uncertainty ($\beta_1+\beta_3$)	0.04

The results reported in table 3 indicate rejection of both hypotheses H_{0_1} and H_{0_2} in favor of the alternative hypotheses. Over the six-year period these data suggest that greater ex *ante* uncertainty in earnings results in lower response coefficients. In table 3, the null hypothesis that $\beta_2=\beta_3=0$ is rejected at the 0.0001 level. Furthermore, β_3 is also significantly more negative than β_2, with H_{0_2} rejected at the 0.0001 level. These results suggest that greater levels of *ex ante* earnings uncertainty result in lower earnings response coefficients.[9]

We repeated these tests on an annual basis to determine whether the results persist in the cross-section each year. The results, reported in table 3, are consistent with the pooled time-series results. The effects of greater uncertainty are revealed by β_2 and β_3, which are negative in 12 of 12 cases. Hypothesis H_{0_1} is rejected at the 0.05 level in three of the six years (1980, 1982, and 1983), and at the 0.061 level in 1984. At the same time, a binomial test shows the probability that β_2 and β_3 would be negative in 12 of 12 cases by chance is less than 0.001. Table 3 also reports the results of tests for H_{0_2}: $\beta_2-\beta_3=0$. The alternative hypothesis stated earlier suggested β_3 should be more negative than β_2. We find β_3 is more negative than β_2 in five of the six years, and the difference is

---

[9] Using several variations of model (5) based on other specifications of *UE*, constant intercept, and time varying intercept, we consistently reject H_{0_1} in favor of the alternative hypothesis. In the pooled time-series results for portfolios reported in panel A and in the comparable firm-specific results not reported here, the effects of *ex ante* uncertainty appear to be significant, with more *ex ante* uncertainty resulting in less informative earnings announcements.

Table 3

Effects of *Ex Ante* Earnings Uncertainty on the Unexpected Earnings-Unexpected Returns Relation

Pooled Results	*Number of Portfolios*	$UR_{is}=\alpha_1+\alpha_2 D_M+\alpha_3 D_H+\beta_1 UE_{is}+\beta_2 D_M UE_{is}+\beta_3 D_H UE_{is}+\epsilon_{is}$[a]							*White's* χ^2 *(Signif.)*	H$_0$: $\beta_2=\beta_3=0$ *Test Statistic*[b] *(Signif.)*	H$_0$: $\beta_2-\beta_3=0$ *Test Statistic* *(Signif.)*
		α_1	α_2	α_3 *(Test statistics in parentheses)*	β_1	β_2	β_3	*Adjusted* R^2			
1979–84	443	−.00218 **(−2.077)**	.00160 (1.080)	.00207 (1.382)	.76765 **(3.556)**	−.34536 (−1.482)	−.72707 **(−3.383)**	.0884	4.0849 (.7699)	14.7949 **(.0001)**	4.3001 **(.0001)**
Year-by-Year											
1979	60	−.00433 (−1.682)	.00376 (1.030)	.00544 (1.497)	1.10102 **(1.679)**	−.64638 (−.938)	−.94736 (−1.434)	.1291	8.0203 (.2366)	1.8391 (.1688)	1.3303 (.0945)
1980	69	−.00695 **(−2.391)**	.00911 **(2.271)**	.00589 (1.453)	1.04396 (1.589)	−.28454 (−.405)	−.93878 (−1.426)	.1992	6.5232 (.4801)	4.3392 **(.0172)**	2.5925 **(.0059)**
1981	77	.00085 (.334)	.00034 (.095)	.00149 (.414)	.47573 (1.104)	−.04583 (−.095)	−.31646 (−.729)	.1133	5.8254 (.5803)	.9937 (.3753)	1.2270 (.1166)
1982	78	.00360 (1.435)	−.00601 (−1.670)	−.00551 (−1.499)	.73463 **(1.105)**	−.26203 (−.379)	−.75231 (−1.131)	.0870	4.7837 (.6863)	3.8947 **(.0243)**	2.5581 **(.0063)**
1983	84	−.00257 **(−.961)**	−.00049 (−.129)	.00040 **(.105)**	1.69340 **(2.667)**	−1.72790 **(−2.359)**	−1.62832 **(−2.562)**	.0910	6.4050 (.3794)	3.3221 **(.0412)**	−.2718 (.6067)
1984	75	−.00505 **(−2.618)**	.00446 (1.620)	.00607 (2.166)	.55624 **(1.805)**	−.30659 (−.908)	−.53916 **(−1.748)**	.0779	7.8889 (.3425)	2.9148 (.0609)	1.6745 **(.0493)**

[a] UR_{is} = unexpected two-day return for portfolio *i* in period (−1, 0);

UE_{is} = unexpected earnings for portfolio *i* in period *s*, measured as the difference between actual earnings and the mean I/B/E/S forecast deflated by stock price two trading days before the earnings announcement;

D_H = 1 for portfolios of firms ranked in the high *ex ante* earnings uncertainty (ASD) group,
= 0 otherwise; and

D_M = 1 for portfolios of firms ranked in the middle *ex ante* earnings uncertainty (ASD) group,
= 0 otherwise.

[b] The test statistic for this hypothesis is a χ^2-statistic (*F*-statistic) when White's test indicates the presence (absence) of heteroskedasticity. The test statistics for the other hypotheses are *z*-statistics (*t*-statistics) when White's test indicates the presence (absence) of heteroskedasticity. One-tail significance at the $p \leq 0.05$ level is indicated by bold face.

significant at the 0.05 level in three of the six years (1980, 1982, and 1984). These results support the contention that the observed differences in response coefficients are related to *ex ante* uncertainty.[10]

The distributional properties of the underlying data were examined to understand why *ex ante* uncertainty might be more significant in some years than in others. This examination revealed that the overall amount of *ex post* uncertainty reflected by the standard deviation of *UE* (reported below) was considerably smaller in 1979 and 1981, two years in which we were unable to reject the hypothesis that $\beta_2=\beta_3=0$. While not conclusive, these results suggest that:

Year	*Standard Deviation of Unexpected Earnings*	H_0: $\beta_2=\beta_3=0$ *from Table 3* *(Significance)*
1979	0.0204	0.1688
1980	0.0398	0.0172
1981	0.0294	0.3752
1982	0.0560	0.0248
1983	0.0599	0.0412
1984	0.0786	0.0609

in years when earnings news (*UE*) is relatively heterogeneous (i.e., larger σ_{UE}), *ex ante* uncertainty is relatively important. Conversely, when earnings news is relatively homogeneous, the importance of *ex ante* uncertainty is reduced.

Uncertainty Diagnostics

Although the results in table 3 support the research hypothesis, it is possible that *ex ante* uncertainty is simply a proxy for other variables that have been demonstrated in prior research to affect the *UR-UE* relationship. Two such variables of interest are firm size and quantity of information available about a firm (Atiase 1985; Freeman 1987). The number of analysts' forecasts in the month preceding the earnings announcement month were used as a measure of the quantity of information available about the firm. More (fewer) analysts' forecasts indicate more (less) interest in the firm and more (less) information about the firm for investors, thus providing a measurable proxy for the amount of information available. Size was measured by market value of common equity of the firm. Both of these variables had Spearman correlation coefficients with *ex ante* uncertainty that were significantly greater than zero and significantly less than one. However, when these variables were added to the equation (5) regression, they were not significant, which suggests that neither firm size nor number of analysts are driving our results.

---

[10] These annual tests were also conducted with firm-specific data rather than portfolios. In comparison to the portfolio results reported in table 3, the firm-specific year-by-year results were somewhat weaker. Yet both sets of alternative results were similar to the portfolio results reported. The β_2 and β_3 coefficients were negative in 12 of 12 cases for the regressions using firm-specific data by year. The probability of obtaining these results by chance is very low ($p<0.001$).

Controlling for Uncertainty

To investigate whether *ex ante* uncertainty in earnings can be controlled in the *UR-UE* relation, we estimate a modified version of equation (5), equation (6), which incorporates *ex ante* uncertainty (analysts' standard deviation) in the measurement of unexpected earnings ($\widehat{UE}$). This alternative specification, in effect, treats ASD as a continuous variable (rather than as a classificatory variable) by dividing *UE* by ASD. If this standardization process controls for *ex ante* uncertainty, we should be unable to reject the corresponding null hypotheses for the following:[11]

$$UR_{it}=\phi_1+\phi_2 D_M+\phi_3 D_H+\gamma_1 \widehat{UE}_{it}+\gamma_2 D_M \widehat{UE}_{it}+\gamma_3 D_H \widehat{UE}_{it}+\epsilon_{it}, \tag{6}$$

where $\widehat{UE}=UE/ASD$, where all other variables are as defined earlier, and

H_{0_1}: $\gamma_2=\gamma_3=0$, and
H_{0_2}: $\gamma_2-\gamma_3=0$.

The results of estimating equation (6), which are reported in table 4, confirm the ability of this standardized measure to reduce differences in response coefficients that are related to differences in ex *ante* uncertainty. The pooled results reported in table 4 for the two-day announcement period $(-1, 0)$ fail to reject the null hypotheses $\gamma_2=\gamma_3=0$, and $\gamma_2-\gamma_3=0$. The γ_2 and γ_3 coefficients in the pooled regression are not significantly different from zero, suggesting that scaling unexpected earnings by ASD controls for the effect of differences in ex *ante* uncertainty to a large extent.

The year-by-year results reported in table 4 also suggest the effects of uncertainty have been controlled for, with γ_2 and γ_3 negative in five of 12 cases. Only in 1983 were both γ_2 and γ_3 negative. Additionally, in no year could the null hypotheses that $\gamma_2=\gamma_3=0$, or that $\gamma_2-\gamma_3=0$ be rejected. In summary, the results from equation (6) reported in table 4 suggest that unexpected earnings as measured by $\widehat{UE}$ appears to control for the differential effects of ex *ante* uncertainty.

IV. Implications and Conclusions

Prior research examining the information content of unexpected earnings has reported results using the *means* or expected values for earnings forecasts in measures of unexpected earnings. Our results suggest that ex *ante* uncertainty in earnings has a systematic effect on the relation between unexpected returns and unexpected earnings. Firms with relatively more (less) ex *ante* uncertainty in earnings appeared to have smaller (larger), less significant (more significant) earnings response coefficients. Theory suggests that uncertainty should have this observed effect on earnings response coefficients when it reflects noise in the earnings signal. Hence, we infer that the dispersion in preannouncement analysts' earnings forecasts is a proxy for noise in the firm's earnings number. When the noise (or garbling) in earnings news is relatively great, the stock price response per unit of earnings news will be relatively small. It is

---

[11] Note that scaling unexpected earnings ($A-F$) by ASD will tend to increase (decrease) the response coefficient most for high (low) ASD firms, since it reduces (increases) the *relative* magnitude of unexpected earnings. Firm-year observations with ASD values of zero were deleted, resulting in 3,099 observations being used for these tests.

Table 4

Effects of *Ex Ante* Earnings Uncertainty on the Unexpected Earnings-Unexpected Returns Relation After Controlling for *Ex Ante* Uncertainty

Pooled Results	Number of Portfolios	$UR_{it}=\phi_1+\phi_2 D_H+\phi_3 D_M+\gamma_1\widehat{UE}_{it}+\gamma_2 D_H\widehat{UE}_{it}+\gamma_3 D_M\widehat{UE}_{it}+\epsilon_{it}$[a]							White's χ^2 (Signif.)	$H_0: \gamma_2=\gamma_3=0$ Test Statistic[b] (Signif.)	$H_0: \gamma_2-\gamma_3=0$ Test Statistic (Signif.)
		ϕ_1	ϕ_2	ϕ_3 (*Test statistics in parentheses*)	γ_1	γ_2	γ_3	Adjusted R^2			
1979–84	434	−.00224 (−1.939)	.00173 (1.079)	.00214 (1.313)	.00133 **(3.755)**	.00076 (1.367)	−.00001 (−.019)	.0910	7.5445 (.4792)	1.1400 (.3208)	1.2947 (.0981)
Year-by-Year											
1979	57	−.00427 (−1.449)	.00390 (.971)	.00545 (1.366)	.00194 **(2.166)**	−.00005 (−.035)	.00113 (0.671)	.1491	9.6257 (.2923)	.2689 (.7653)	−.6817 (.7507)
1980	66	−.00846 **(−2.737)**	.01040 **(2.506)**	.00730 **(1.754)**	.00252 **(2.204)**	.00176 (1.083)	.00440 **(2.341)**	.3806	8.0325 (.4303)	2.7435 (.0724)	−1.4037 (.9172)
1981	76	.00116 (.463)	.00018 (.052)	.00041 (.117)	.00133 (1.579)	.00075 (.520)	.00278 **(1.760)**	.1224	7.3922 (.4950)	1.5500 (.2195)	−1.1433 (.8716)
1982	77	.00351 (1.092)	−.00591 (−1.294)	−.00426 (−.896)	.00097 (.796)	.00124 (.701)	−.00050 (−.256)	.0165	8.4523 (.3906)	.4349 (.6490)	.8723 (.1930)
1983	83	−.00276 (−1.043)	−.00022 (−.058)	−.00061 (−.162)	.00216 **(2.810)**	−.00173 (−.924)	−.00155 (−1.261)	.0423	7.4852 (.4853)	.9872 **(.3773)**	−.0954 (.5379)
1984	75	−.00505 **(−2.335)**	.00465 (1.512)	.00588 **(1.856)**	.00051 (.986)	.00096 (1.230)	−.00028 (−.386)	.0695	11.5299 (.1734)	1.3741 (.2599)	2.6061 (.0555)

[a] UR_{it} = unexpected two-day return for portfolio i in period (−1, 0);

$\widehat{UE}_{it}$ = unexpected earnings for portfolio i measured as the difference between actual earnings and the mean I/B/E/S forecast deflated by the standard deviation of analysts' forecasts;

D_H = 1 for portfolios of firms ranked in the high *ex ante* earnings uncertainty (ASD) group,
= 0 otherwise; and

D_M = 1 for portfolios of firms ranked in the middle *ex ante* earnings uncertainty (ASD) group,
= 0 otherwise.

[b] The test statistic for this hypothesis is a χ^2-statistic (F-statistic) when White's test indicates the presence (absence) of heteroskedasticity. Significance at the $p \leq 0.05$ level is indicated by bold face. The test statistics for the other hypotheses are z-statistics (t-statistics) when heteroskedasticity is (is not) significant. One-tail significance at the $p \leq 0.05$ level is indicated by bold face.

possible that greater *ex ante* earnings uncertainty is a signal of relatively lower earnings quality. Could uncertainty signal earnings management? Future research may provide further evidence on the relation between earnings quality, earnings management, and *ex ante* uncertainty.

In addition to documenting the effect of *ex ante* uncertainty on the unexpected earnings-unexpected returns relation, we provide a measure of unexpected earnings that to a large extent controls for cross-sectional differences in *ex ante* uncertainty. This measure of unexpected earnings (which reflects both the mean and the variance of earnings predictions) may prove useful in future cross-sectional studies concerned with the information content of unexpected earnings. It may be particularly appropriate in multivariate tests in which controlling for the effects of earnings uncertainty is important for interpreting the information content of other (nonearnings) variables.

References

Atiase, R. 1985. Predisclosure information, firm capitalization and security price behavior around earnings announcements. *Journal of Accounting Research* 23 (Spring): 21–35.

——. 1987. Market implications of predisclosure information: Size and exchange effects. *Journal of Accounting Research* 25 (Spring): 168–76.

Bernard, V. 1987. Cross-sectional dependence and problems in inference in market based accounting research. *Journal of Accounting Research* 25 (Spring): 1–48.

Christie, A. 1987. On cross-sectional analysis in accounting research. *Journal of Accounting & Economics* 9 (December): 231–58.

Cornell, B., and W. R. Landsman. 1989. Security price response to quarterly earnings announcements and analysts' forecast revisions. *The Accounting Review* 64 (October): 680–92.

Easton, P., and M. E. Zmijewski. 1989. Cross-sectional variation in the stock market response to earnings announcements. *Journal of Accounting & Economics* 11 (July): 117–42.

Epstein, L. G., and S. M. Turnbull. 1980. Capital asset prices and the temporal resolution of uncertainty. *Journal of Finance* 35 (June): 627–43.

Freeman, R. N. 1987. The association between accounting earnings and security returns for large and small firms. *Journal of Accounting & Economics* 9 (July): 195–228.

Grant, E. 1980. Market implications of differential amounts of interim information. *Journal of Accounting Research* 18 (Spring): 255–68.

Holthausen, R. W., and R. E. Verrecchia. 1988. The effect of sequential information releases on the variance of price changes in an intertemporal multi-asset market. *Journal of Accounting Research* 26 (Spring): 82–106.

Imhoff, E., and G. Lobo. 1991. Characteristics of earnings forecast errors: Some empirical evidence. Working paper, Syracuse University, NY.

Kmenta, J. 1971. *Elements of Econometrics*. New York: Macmillan.

Kormendi, R., and R. Lipe. 1987. Earnings innovations, earnings persistence, and stock returns. *Journal of Business* 60 (July): 323–45.

Lang, M. 1989. Time-varying stock price response to earnings induced by uncertainty about the time-series process of earnings. Working paper, University of Chicago.

Lipe, R. 1990. The relation between stock returns and accounting earnings given alternative information. *The Accounting Review* 65 (January): 49–71.

Pincus, M. 1983. Information characteristics of earnings announcements and stock market behavior. *Journal of Accounting Research* 21 (Spring): 155–83.

Rao, G. R. 1989. The relation between stock returns and earnings: A study of newly-public firms. Working paper, University of Rochester, NY.

Robichek, A. A., and S. C. Myers. 1966. Valuation of the firm: Effects of uncertainty in a market context. *Journal of Finance* 21 (June): 215–27.

Verrecchia, R. E. 1980. The rapidity of price adjustments to information. *Journal of Accounting & Economics* 2 (March): 63–92.

White, H. 1980. A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity. *Econometrica* 48 (May): 817–38.

THE ACCOUNTING REVIEW
Vol. 67, No. 2
April 1992
pp. 440–442

COMMENTS

A Comment on "The Valuation of R&D Firms with R&D Limited Partnerships"

Joseph K. Cheung
George Mason University
Mandy Li
University of Maryland

IN a recent article, Shevlin (1991) models an R&D limited partnership (LP) as a call option held by the R&D firm (the LP call option) and shows that the value of this call option is reflected in the R&D firm's stock price. His overall conclusion is that investors find existing footnote disclosures useful in valuing the LP to the R&D firm. More specifically, he also suggests that the underlying-asset and exercise-price components of the LP call option estimate the assets and liabilities, respectively, of the LP.[1] This comment aims to clarify his interpretation of the components of the LP call option.

In the LP arrangements studied by Shevlin, the R&D firm is required to pay a royalty to the LP if it chooses to exploit the technology developed by the LP. Shevlin models this type of arrangements as a Margrabe (1978) call option, a generalized version of the Black and Scholes (1973) formula in that the option's exercise price is stochastic rather than constant. To avoid cluttering, we will discuss the before-tax version of Shevlin's equation 2′. Following Shevlin, let P and X be the current values of, respectively, the R&D outcome and the royalty payment. Before taxes, the LP call option has a value given by:

$$PN(d_1)-XN(d_2), \tag{1}$$

where $N(d_1)$ and $N(d_2)$ are as defined in his equation 2. Shevlin interprets $PN(d_1)$ and $XN(d_2)$ as estimates of the "asset and liability components of the LP" (p. 6), thus

---

[1] In the finance literature, the term "liabilities" is used to refer to all the claims—debt and equity—a corporation has outstanding. For instance, Black and Scholes (1973, 648–49) include as "corporate liabilities" common stock, bonds, and warrants. The term "debt" as used by finance writers is closer to being the accounting concept of liability. In Shevlin (1991) as in this comment, the term "liabilities" is used in the accounting sense.

Submitted March 1991.
Accepted October 1991.

"increasing both the assets and the liabilities of the R&D firm" (p. 17). That is, based on the analogy of corporate equity to a call option, Shevlin views the LP call option as an equity interest in the LP and seeks to determine the asset and liability components of the equity using an option formula.

We observe that $PN(d_1)$ and $XN(d_2)$ are incorrect formulations of these asset and liability components in the option pricing framework. It is true that, according to Black and Scholes (1973, 649), a firm's equity can be viewed as a call option on the firm's assets, with an exercise price equal to the maturity value of the firm's debt. By this, however, Black and Scholes merely indicate how the firm's equity can be valued; they do *not* also suggest that the firm's assets and liabilities can be valued as the underlying-asset and the exercise-price components of the call option. In fact, they (p. 650) explicitly indicate how the firm's liabilities are to be priced—as the difference in value between the firm's assets and equity, the latter modeled as a call option.

This can be stated precisely. Because of the symmetry of the normal distribution, we have the following identity based on equation (1):

$$PN(d_1)-XN(d_2)\equiv P-\{X-[XN(-d_2)-PN(-d_1)]\}. \tag{2}$$

We agree that the LP's equity can be valued as the left-hand side of equation (2). However, the asset and liability components of the LP should be valued by, respectively, the first and second term on the right-hand side of equation (2). The first term P on the right-hand side is indisputably the LP's assets, given Shevlin's definition (p. 5) and measurement (p. 10) of the variable. The second term $\{X-[XN(-d_2)-PN(-d_1)]\}$ consists of a default-free value of the LP's liabilities, X, minus a default premium in the form of a put option. Thus, the second term is consistent with the notion of a risky debt. Viewed differently, the second term is equal to P, the LP's asset value, minus the left-hand side, the LP's equity value, which is exactly what Black and Scholes have intended as the valuation of the LP's liabilities in an option pricing framework.[2]

Obviously, unless $N(d_1)=1$, the *corresponding* terms on the two sides of equation (2) are not equal. Thus, $PN(d_1)$ and $XN(d_2)$ are incorrect measures of, and have economic interpretations different from, the LP's assets and liabilities. In particular, the term $N(d_2)$ in the exercise-price component is the market value of a security that pays \$1 at time T if $\tilde{P}_T>\tilde{X}_T$ and zero otherwise (Breeden and Litzenberger 1978, 630). This means that $-XN(d_2)$ is a state-contingent outflow, effectively a short position in a portfolio of Arrow-Debreu securities. It is debatable, therefore, that $-XN(d_2)$ fits into the usual notion of liability. It is true that, under the assumption of risk neutrality, $PN(d_1)$ is a present value of the conditional expectation of $\tilde{P}_T$, and $XN(d_2)$ is a present

---

[2] One might question why the Black-Scholes notions of corporate asset, liability, and equity apply in the Margrabe model. The reason can be stated formally as follows. Applying Stulz's (1982) equation 11′ to equation (1), we have:

$$PN(d_1)-XN(d_2)\equiv P-M(P,X,O,T),$$

where $M(P,X,O,T)$ is the price of a European call option on min (P,X), with a zero exercise price and time to maturity T. Stulz shows that the maturity value of $M(P,X,O,T)$ is $\max\{\min(\tilde{P}_T,\tilde{X}_T)-0,0\}$, which is simply $\min(\tilde{P}_T,\tilde{X}_T)$, where $\tilde{P}_T$ and $\tilde{X}_T$ are maturity values. But $\min(\tilde{P}_T,\tilde{X}_T)$ corresponds to the payoff on a corporate debt with maturity value $\tilde{X}_T$. Hence, the right-hand side of the equation above, and thus the right-hand side of equation (2), measures the asset and liability components of a firm whose equity is formulated as a Margrabe option.

value of the expected exercise price, since $N(d_2)=\text{Prob}(\tilde{P}_T>\tilde{X}_T)$ in a risk-neutral economy (Cox et al. 1979, 235–36). In fact, it seems that Shevlin adopts this risk-neutral interpretation by referring to the components of the option formula as "present values of the expected inflows and outflows of the R&D firm arising from the LP" (p. 4). However, the risk-neutral interpretation of the option formula is purely fictional. Thus, $PN(d_1)$ and $XN(d_2)$ do not in general measure the LP's expected assets and liabilities.

Although this observation is unlikely to affect Shevlin's overall conclusion, it does have a bearing on his analysis of the LP's assets and liabilities. For instance, in regard to his equation 6, the variables *LPCA* and *LPCD*, which are after-tax versions of $PN(d_1)$ and $XN(d_2)$, are biased measures of the R&D firm's assets and liabilities arising from the LP relationship. Consequently, it is not warranted to expect the regression coefficients of *LPCA* and *LPCD* to be 1 and –1, respectively (p. 12), given that the same values are expected of the regression coefficients of the R&D firm's reported assets and liabilities.

In summary, we do not dispute that Shevlin's analysis supports the overall conclusion that investors find existing footnote disclosures useful in valuing the LP to the R&D firm. We only disagree with how the asset and liability components of the LP are measured in the option pricing framework and offer a correction accordingly.

References

Black, F., and M. Scholes. 1973. The pricing of options and corporate liabilities. *Journal of Political Economy* 81 (May/June): 637–54.

Breeden, D., and R. Litzenberger. 1978. Prices of state-contingent claims implicit in option prices. *The Journal of Business* 51 (October): 621–51.

Cox, J., S. Ross, and M. Rubinstein. 1979. Option pricing: A simplified approach. *Journal of Financial Economics* 7 (September): 229–63.

Margrabe, W. 1978. The value of an option to exchange one asset for another. *The Journal of Finance* 33 (March): 177–86.

Shevlin, T. 1991. The valuation of R&D firms with R&D limited partnerships. *The Accounting Review* 66 (January): 1–21.

Stulz, R. 1982. Options on the minimum or the maximum of two risky assets: Analysis and applications. *Journal of Financial Economics* 10 (July): 161–85.

THE ACCOUNTING REVIEW
Vol. 67, No. 2
April 1992
pp. 443–445

A Reply to "A Comment on 'The Valuation of R&D Firms with R&D Limited Partnerships'"

Terry Shevlin[1]
University of Washington

THE comment by Cheung and Li aims to clarify my interpretation (Shevlin 1991) of the components of the call option held by the R&D firm with respect to its investment in an R&D limited partnership. As noted by Cheung and Li, I value the R&D firm's interest in the limited partnership as a call option (*LPCO*) using the Margrabe (1978) model:[2]

$$LPCO = PN(d_1) - XN(d_2), \tag{1}$$

where P is the estimated present value of the R&D project funded by the limited partnership, X is the estimated present value of the exercise price (a random variable), $N(\bullet)$ is the cumulative standard normal distribution function, and d_1 and d_2 are defined in Shevlin (p. 5). I interpret $PN(d_1)$ as the value of the asset (*LPCA*) and $XN(d_2)$ as the liability (*LPCD*) components of the call option held by the R&D firm (p. 6). Cheung and Li argue that this interpretation is incorrect. Following Black and Scholes (1973, 650), Cheung and Li show that equation (1) above can be restated as (from their eq. [2]):

$$LPCO = P - (X - [XN(-d_2) - PN(-d_1)]). \tag{2}$$

Note that equation (2) is simply an algebraic manipulation of equation (1) (using $gN(d_1) \equiv g - gN(-d_1)$). The comment then argues that:

$$LPCA = P,$$

and

$$\begin{aligned} LPCD &= X - [XN(-d_2) - PN(-d_1)], \\ &= XN(d_2) + PN(-d_1), \text{ and} \\ &= P - LPCO. \end{aligned}$$

---

[1] The helpful comments of Thomas Hemmer and D. Shores are appreciated. Shevlin is visiting at the University of Iowa for the 1991–92 academic year.

[2] Consistent with Cheung and Li, I use the pre-tax version here. In the Margrabe model, P and X are the market values of two risky assets. In my setting, there is no market value for either asset and I have to estimate them using present values.

Received December 1991.

Table 1

Cross-sectional Regressions of Market Value of Equity on Total Reported Assets, Total Reported Debt, In-house R&D Expenditures, Estimated Value of the LP Funded R&D, and the Estimated Values of the Asset and Liability Components of the LP Call Option

$$S_j=\alpha_0+\alpha_1 A_j+\alpha_2 D_j+\alpha_3 IHRD_j+\alpha_4 LPCA_j+\alpha_5 LPCD_j+u_j$$

Predicted Sign	α_0	α_1 +	α_2 −	α_3 +	α_4 +	α_5 −	*Adjusted* R^2
1. Shevlin (1991)[a]	7.050	2.544	−3.887	1.058	0.840	−1.884	.837
t-statistic ($\alpha=0$)[b]	3.896	11.886	−9.366	5.365	10.492	−5.185	
t-statistic ($\alpha=1$)		7.214	−6.956	0.296	−2.004	−2.433	
2. Cheung and Li (1992)[c]	6.798	2.316	−3.648	1.216	0.644	−0.844	.832
t-statistic ($\alpha=0$)	3.686	11.710	−8.941	5.907	11.356	−8.940	
t-statistic ($\alpha=1$)		6.653	−6.489	1.048	−6.271	1.652	

[a] Results reported in Shevlin (1991), Table 4, regression (1). Estimation based on 53 sample firms with 145 yearly observations. All variables are deflated by the book value of equity. *S* is the market value of shareholders' equity, *A* is the book value of total assets, *D* is the estimated market value of debt, *IHRD* is k_t* in-house R&D expenditures, where k_t is an estimated market capitalization factor in period t, $LPCA=PN(d_1)$ is an estimate of the asset component of the LP call option, and $LPCD=XN(d_2)$ is an estimate of the liability component of the LP call option.

[b] The first row of t-statistics tests the null hypothesis that the $\alpha_i=0$ for $i=0$ to 5. The second row of *t*-statistics tests the null hypothesis that $\alpha_i=1$ for $i=1$, 3, and 4, $\alpha_i=-1$ for $i=2$ and 5.

[c] *LPCA* estimated as *P* and *LPCD* estimated as $XN(d_2)+PN(-d_1)=P-LPCO$ as per suggestion in the comment (Cheung and Li 1992). Note that in both regressions the option variables are stated on an after-tax basis.

The different formulations for *LPCA* and *LPCD* arise because of differing economic interpretations. My formulation arises because, as noted by Cheung and Li, I implicitly interpreted $XN(d_2)$ as a state-contingent outflow and $PN(d_1)$ as a state-contingent inflow, with the contingency being that the call option finish in-the-money. Cheung and Li state that it is debatable whether this interpretation fits the usual notion of a liability. I continue to believe my interpretation fits the usual notion of a contingent liability but the important point is whether my decomposition of the call option into asset and liability components represents an economic asset and liability rather than current GAAP definitions of assets and liabilities.

The difference between my formulation and that by Cheung and Li basically deals with how investors construct the economic balance sheet of the firm. Do investors include $PN(d_1)$ or *P* on the left-hand side and $XN(d_2)$ or $XN(d_2)+PN(-d_1)$ on the right-hand side of the economic balance sheet.[3] Given the formulation in equation (2), Cheung and Li go on to argue that my estimates of *LPCA* and *LPCD* contain measurement error and it is, therefore, not surprising that the estimated coefficients differ from their predicted values of +1 and −1, respectively. However, my estimates will not contain measurement error (ignoring other sources of measurement error in the empirical proxies) if investors construct their economic balance using equation (1). It is an empirical question how investors construct their balance sheets.

---

[3] Two analogies with the accounting balance sheet provide some intuition for the debate here. Should interim payments received on long term-contracts be classified as a liability (unearned revenue) or as a contra asset account (offsetting inventory)? Should investment in long-term bonds purchased at a discount be reported as an asset at its present value of future cash flows or at face value with the discount reported as a liability (basically, unearned interest)? The net amount to be reported is the same in both instances, but the amount to be reported on the asset and liability sides of the accounting balance sheet differs among the alternatives in each example. I believe current GAAP treatment of these two items is consistent with my formulation.

I reestimated the regression model (eq. [6] in Shevlin 1991) using the formulation of *LPCA* and *LPCD* suggested in Cheung and Li. The results are reported in table 1. The estimated coefficient on *LPCA* decreases from 0.840 to 0.644 and continues to be significantly less than its predicted value of +1 (and continues to be significantly greater than zero). The estimated coefficient on *LPCD* increases from −1.884 to −0.844 and is not significantly different from −1 at conventional levels in a two-tailed test. There is a slight decline in the adjusted R^2. Overall, I believe these results suggest it is difficult to conclusively argue that either interpretation is superior to the other. Nevertheless, as noted by Cheung and Li, the original conclusions of Shevlin (1991) still hold: the evidence is consistent with equity market prices reflecting the information (disclosed in the notes to the accounts) about R&D firms interests in R&D limited partnerships.

References

Black, F., and M. Scholes. 1973. The pricing of options and corporate liabilities. *Journal of Political Economy* (May/June): 637–54.

Cheung, J. K., and M. Li. 1992. A comment on "The valuation of R&D firms with R&D limited partnerships." *The Accounting Review* 67 (April): 440–42.

Margrabe, W. 1978. The value of an option to exchange one asset for another. *The Journal of Finance* (March): 177–86.

Shevlin, T. 1991. The valuation of R&D firms with R&D limited partnerships. *The Accounting Review* 66 (January): 1–21.

THE ACCOUNTING REVIEW
Vol. 67, No. 2
April 1992
pp. 446–456

BOOK REVIEWS

Shane R. Moriarity, Editor

Editor's Note: Two copies of books for review should be sent to Professor Shane R. Moriarity, School of Accounting, University of Oklahoma, Norman, OK 73019. The policy of the *Review* is to publish only those reviews solicited by the Book Review Editor. Unsolicited reviews will not be accepted.

DONALD A. CURTIS, *Management Rediscovered: How Companies Can Escape the Numbers Trap* (Homewood, IL: Richard D. Irwin, 1990, pp. xi, 200, $24.95).

Donald A. Curtis, management consultant, postulates that many of the basic flows and problems found in the American management system are based on two fundamental premises:

> The important characteristics of a business can be quantified. This ability to quantify can be used to directly support all the management activities (i.e., understanding, directing, and controlling the enterprise) (p. 7).

The author believes that these two premises are false and that they lead to a management system he calls "Analogue Management" of managing through quantitative representations of reality. The book is divided into two parts. In the first part the author discusses in length the shortcomings of analogue management and indicates its pitfalls. The second part of the book is filled with suggestions on what managers can do to escape the current system and thereby do more than "manage by numbers."

In Part I Curtis contends that American companies fall into the so-called "analogue management trap" when they believe that the analogue measurements of business are more comprehensive and more accurate than they are in reality. Our most widely used analogue, the American accounting system, is described as a badly flawed analogue, essentially useless as a measurement of overall corporate performance in the time periods in which it is used. In particular, the accounting system is criticized for being too financial-transaction oriented and biased toward short-run behavior. The result is to bias accounting asset valuation away from off-balance-sheet assets such as the quality of the human organization and the company's market position. Since accounting measurements are not a true measure of corporate performance, this has created an environment which has led to a great deal of counterproductive management behavior that has resulted in the waste of a large amount of corporate resources.

With the basic framework established in Part I, the author devotes Part II to suggesting ways that companies can avoid the analogue management trap. Unfortunately the author does not have a completely clear-cut path for one to follow to the Utopian land of reality. The book will be a disappointment to readers who expect a "quantum leap" out of the analogue management trap. In fact, the author points out that none of his suggestions require a company to abandon any of its current reporting or management practices, although changes in both are clearly recommended. Instead, he suggests ways to modify the current system and thereby create more relevant and useful data for management. The end result is actually a plan for more analogues. For example, he suggests the use of various analogues he considers to be more objective and accurate than such abstracts as monthly accounting reports. Overall company performance would include measures of product value to customers by means of defects in the first year of use, overall productivity by means of sales per employee, and the use of customer service indices such as periodic customer satisfaction surveys.

Additionally, he suggests creating analogues to help management better understand and control the more detailed aspects of the business, such as time required to process customer orders, errors per 1,000 invoices, and employee turnover.

Later chapters in the book deal with methods to make the management reports right technically. A distinction is drawn between reports used for controlling and those used for measuring the various aspects of the business. He describes some principles—e.g., responsibility accounting and flexible budgeting—that must be followed if the reports are to be correct.

The final stages of the book move away from micromanagement techniques to a discussion of a company's overall strategy. According to the author the basis of competition is ultimately an organizational issue and, in this context, the author introduces a concept called "continuous competitive advantage." Continuous competitive management focuses on maintaining a mainline or "core" business and on enhancing the company's human resources (off-balance-sheet assets) through the acquistion, development, and retention of people in the organization.

Basically, the book contains no new ideas about managing organizations. But this should not suggest that the book does not make a worthwhile contribution. The author does an excellent job of making the reader aware of the limitations of our analogue-based management system and demonstrates what managing through this system does to both the organization and to its individual managers. Overall, this book is well-conceived and well-written. It is an enjoyable book to read and it should appeal to a wide variety of readers in the academic and business world.

FRED A. THORNTON, JR.
Associate Professor of Accounting
University of West Florida

DALE FLESHER, *The Institute of Internal Auditors: 50 Years of Progress Through Sharing* (Altamonte Springs, FL: The Institute of Internal Auditors, 1991, pp. xi, 180, $15.00 members, $20.00 nonmembers).

The book begins by providing a brief history of internal auditing before it was organized into a profession. The author then recounts how a committee of three men—Victor Brink, John Thurston, and Robert Milne—formed The Institute of Internal Auditors (IIA) in 1941. Although Thurston and Milne had been contemplating an independent organization for internal auditors, it was the publication of Brink's doctoral dissertation that provided the spark to get these men together and form the IIA. (How often does a doctoral dissertation in accounting make this kind of impact!) The remainder of the first chapter focuses primarily on the evolution of operational auditing. The author discusses how the U.S. General Accounting Office pioneered operational auditing in the United States and how operational auditing has developed in several other countries.

The following two chapters discuss events that have enhanced the internal auditing profession. Some of these were initiated by the IIA (e.g., a code of ethics, standards, certification), while others occurred externally (e.g., Foreign Corrupt Practices Act, Treadway Commission). Particularly interesting were the discussions about the "grandfathering" provisions for the CIA certificate and the controversy surrounding the CIA acronym.

The next three chapters revolve around various types of professional services that have been offered by the IIA. After a brief discussion about the acquisition of membership, the author describes the evolution of IIA chapters and affiliated national institutes. The international representation, in particular, is quite impressive. The book then provides historical perspectives on *Internal Auditor* (the journal published by the IIA) and the various newsletters the IIA has had. Next, the author describes the involvement of the IIA with higher education. Specifically, the IIA's support (or, at times, lack thereof) of internal audit courses, internal audit programs, doctoral students, and educator awards are discussed.

The two chapters that follow are devoted to the professional development activities of the IIA. These include international and regional conferences, educational seminars, and various research activities. The latter consist of internally-conducted research as well as research contracted out to academics.

The final section of the book is devoted to the IIA's past leaders and award winners. Capsule biographies are provided about numerous individuals who held leadership positions and/or who won major IIA awards.

Overall, the book is a well-written and well-organized history of the internal auditing profession during the past 50 years. While the book's brevity is one of its greatest strengths, there were several instances where I was curious to know not only what happened, but also why it happened. For instance, why were academics not permitted to sit for the CIA exam some years ago (p. 43)? Or, why was there a major attendance problem for the international conference at Jerusalem in 1983, but not for the international conference at Sydney in 1985 (p. 116)? Naturally, those who have been active in the IIA will be most interested in reading this book. In addition, others who have been involved with internal auditing education or research will also find the book informative and useful.

ARNOLD SCHNEIDER
Associate Professor of Accounting

Georgia Institute of Technology

C. G. A. FLETCHER, *Government Accounting* (Edinburgh, Scotland: The Institute of Chartered Accountants, 1991, pp. vii, 123, £7.50).

This monograph, eleventh in a series sponsored by The Institute of Chartered Accountants of Scotland, is designed to present to the non-specialist reader the basics of central government financial planning, control and reporting, or as the author indicates in the preface, "what is accounted for and in which documents 'it' can be found" (1). If one is to condense such a work into 123 pages of text, few descriptive adjectives will be used and much lustre will be sacrificed.

To guide those of us who are nescient in the specifics of the British language as related to economics and governmental accounting, the author provides a brief glossary covering two and one-half pages, and a listing of 36 acronyms which are used throughout the book. This is followed by a succinct discussion of public sector functions and the accounting documents which link them together. Before going too far, the reader is convinced that governmental accounting systems are not designed as a whole at one time and, in spite of the efforts of concerned individuals, reports can outlive their usefulness.

Perhaps the most interesting part of this work is found in the fourth chapter, which presents a 300 year chronology of major events in the accounting and finance of the British government. While progress is often difficult to detect in central governments, the author paints a clear picture of the successes and, as frequently, the failures of reform measures and experiments. It becomes obvious that those in power are not necessarily attracted to greater accountability or the revealing of sensitive areas. This chapter is amplified by a subsequent discussion (Chap. 9) of the changes which are being initiated currently. The balance of the book is devoted to a somewhat tedious exposition of sundry documents and reports, their overall content, the budgeting process, and cash flows.

There are three particulars of which any prospective reader should be aware. First, the type in the furnished copy lacked a crispness normally attained by most desktop laser printers. Letters had bled and in many words there was no kerning whatsoever. The effect of this poor production is that reading is slowed significantly. As an ancillary to this, two of the three figures in the book were probably prepared using a software package that apparently does not have a Times Roman font, the choice for the rest of the book. Once again letters bled together to make reading even more difficult and detract from understanding. One must wonder why?

Second, the heavy use of acronyms detracts from the text. While some of them may be familiar (e.g., BBC), readers who are not familiar with the subject may have difficulty laboring through the other acronyms. Third, one should be prepared to translate an occasional very British colloquialism such as "gilts," "hived off," "tranches," and "outturns."

The average reader of this journal will find little benefit in the reading of this book. It may have limited appeal to those who are attempting to internationalize their governmental accounting or public finance courses. While the book does achieve its objective of providing a brief overview of the accounting of the central government, it is that brevity which will prevent it from being an authoritative source.

DAVID R. L. GABHART
Professor of Accountancy

Bentley College

O. FINLEY GRAVES, Editor, *The Costing Heritage: Studies in Honor of S. Paul Garner* (Harrisonburg, VA: The Academy of Accounting Historians, 1991, pp. ix, 171, $15.00 individuals, $20.00 institutions).

This collection of papers deals primarily with cost accounting and its history. The preface refers to it as a *festschrift*, from the German, literally festival+writing, with a dictionary definition of "a volume of learned articles or essays by colleagues and admirers, serving as a tribute or memorial especially to a scholar." The term "*festschrift*" is, in fact, appropriately descriptive of this monograph.

The tribute is to S. Paul Garner (Dean and Professor Emeritus of the University of Alabama) on the occasion of his eightieth birthday in 1990. The focus of this collection matches well two of his life-long interests: (1) his work in the cost accounting field, including his cornerstone study, *Evolution of Cost Accounting to 1925* (first published in 1954 by the University of Alabama Press), and (2) his work to develop the international dimension of accounting. It is these two interests that are richly expounded in two essays of tribute to Professor Garner and the ten studies which follow.

This monograph is the sixth in a series published by the Academy of Accounting Historians. With the exception of one of the tribute essays (Chinese author) and two of the articles (one Japanese, one German author), all of the contributors are from English-speaking countries—two U.S., the others from a worldwide geography. While the authors are in the majority from English-speaking countries, their essays present and analyze historical developments from throughout the world.

The periods covered range from an article about archeological research on counting and token accounting (ca. 8000 to 3000 B.C.), to the contemporary events surrounding the reestablishment of the Cost Accounting Standards Board in 1988 and the nature of activity-based costing, including an analysis of recent work by Johnson, Kaplan, and Cooper.

The arrangement of articles appears to be random. Alternatively, a sequencing moving from studies of the oldest historical periods, chronologically to recent developments would have made for a smoother, more logical flow for one who elects to read the entire monograph. This, of course, is not a particular problem to those doing only selective reading.

Selection is especially enhanced by a well-prepared abstract preceding each study. Extensive bibliographies are found at the end of each study, affording a wealth of materials for further study. Numerous references are made to Garner's *Evolution of Cost Accounting to 1925*, which in itself so carefully documents our historical roots.

The centuries-old need for relevant management information comes into sharp relief as one probes the details of these essays. Each is well-balanced between being both historically descriptive and analytical. From such studies we see threads that even today form the fabric of information provided for control and for decision making, tempered by the necessary structure required in today's economic environment. Indeed, the past can be instructive.

MILTON F. USRY
Professor of Accountancy
University of West Florida

S. J. GRAY, C. B. ROBERTS, and P. D. GORDON, *Making Corporate Reports Valuable: A Survey of Corporate Reporting Practices by Major U.K. Companies* (Edinburgh: The Institute of Chartered Accountants of Scotland, 1991, pp. Volume 1: 43, Volume 2: 212, £20.00).

In 1988 the Institute of Chartered Accountants of Scotland published *Making Corporate Reports Valuable* (*MCRV*) to stimulate debate and encourage improvement in corporate reporting. One criticism of that Report was that the recommendations were radical and impractical. The two volumes published in 1991 (with the same main title as the 1988 Report) address this criticism by identifying U.K. companies which currently disclose information consistent with the 1988 *MCRV* recommendations. Volume 1 describes a survey of the extent to which a sample of 300 major U. K. firms practice 59 specific disclosure recommendations from *MCRV*. Volume 2 provides (without commentary) 71 examples of innovative disclosures.

A particular strength of the study is the inclusion of both periodic reports (audited and unaudited) and non-periodic reports such as takeover documents. Overall, current practice is a long way from the ideal recommended in *MCRV*. One company displays 31 of the recommended practices in their periodic reports but most companies display between 6 and 11 recommended disclosure practices. The survey results are disaggregated into seven areas: financial position and performance; segment information; innovation and R&D; external and market information; personnel; future-oriented information; and, audit. Within each of these areas, the results are broken down by size of firm (defined by "turnover" [revenue]). The survey of non-periodic reports is of particular interest. One finding worthy of further investigation is that the extent of disclosure during hostile takeovers is more consistent with the *MCRV* recommendations.

This survey is intended to serve a particular polemic purpose, i.e., to convince skeptics of the possibility of implementing *MCRV* recommendations, and, perhaps, should not be evaluated against academic norms. But it is evident that the approach taken in this survey and the earlier report are not anchored in contemporary research. The *MCRV* recommendations are presented as a "one-size-fits-all" package but it is likely that different firms will find some recommendations more relevant than others. The finding of significant variation by firm size is one indication of the need to consider contingencies. The creation of a checklist of disclosure practices avoids questioning the relative informativeness of disclosures. For example, the repositioning of the auditor's report is not as informative as the disclosure of segment information.

These volumes provide an interesting snapshot of voluntary disclosures in the U.K. but are likely to have a limited audience in North America. The second volume provides a wealth of examples of innovative disclosures which might be useful in intermediate accounting courses.

ALAN J. RICHARDSON
Associate Professor

Queen's University, Canada

DAVID J. HARR and JAMES T. GODFREY, *Private Sector Performance Measures and Their Applicability to Government Operations* (Montvale, NJ: National Association of Accountants, 1991, pp. xi, 93, $14.95).

This research study, which is jointly sponsored by the National Association of Accountants and the Association of Government Accountants, provides additional insights into the topic of measuring government service efforts and accomplishments (SEA). Unlike the recent series of SEA research studies focusing on state and local government published by the Government Accounting Standards Board, these authors investigate cost measures used in private corporations and compare those with measures currently used by a Federal government agency (Defense Logistics Agency). This detailed investigation of a government agency provides a valuable reference for those interested in the operations of government and how accounting measures might impact that operation.

The intended audience for this study is practicing managerial accountants, particularly those involved in public sector organizations. The book includes a section on selected basic cost accounting terminology and methodology and then proceeds to describe the results of a survey and field study to ascertain cost management techniques and financial performance measures of private sector firms' logistic operations. The usable sample consisted of 22 survey responses and 10 field visits. The field study results provide insights into current methods used by the private sector for cost management. These are then used as a baseline to compare with public sector cost management practices.

The most significant contribution of the book to those interested in public sector management is the section which describes the cost management techniques and financial performance measures used by the Defense Logistics Agency (DLA). The authors give a detailed description of the implementation of new financial performance measures. Most important is the description of how the performance measures are used as a motivational device to contain costs. I suspect that the described bonus plan used by the DLA to reward employees for effective cost management is atypical, but could possibly be used as a model by other public sector agencies providing similar services.

The authors offer suggestions to accountants on how they should become involved in cost management in both public and private sector organizations. This "challenge" is an attempt to motivate accountants to broaden their participation in the internal management of both public and private sector organizations, thereby demonstrating their contribution to organizational objectives.

The detailed study of a single Federal government agency provides the most significant contribution of this study, but can also be considered its limitation. It is questionable whether the practices of a selected government agency generalize to other public sector operations. The authors acknowledge that the financial performance measures used in a single agency may not be universally applied to other agencies. The generalizability of methods described in the study would appear to be particularly limited in its application to government agencies whose service objective and achievement is difficult to measure (such as education).

This study contributes to the body of knowledge relating to the measurement of government service efforts and accomplishments and is recommended for those who are interested in a well-written and detailed case study on government cost measurement. Along with other recent theoretical and descriptive research on government SEA measures, our knowledge of the role accounting measures play in government management continues to advance.

MARC A. RUBIN
Assistant Professor of Accountancy
Miami University

TAKASHI KANATSU, *TQC for Accounting: A New Role in Companywide Improvement* (Cambridge: Productivity Press, English translation, 1990 pp. xvii, 222, $45.00).

This book is not concerned with applying total quality control (TQC) to accounting nor is it concerned with accounting's involvement in company-wide quality control (QC) efforts. Mr. Kanatsu, a Japanese management consultant, previously authored *A TQC Guide to Sales*. Mr. Kanatsu's primary concern in *TQC for Accounting* is getting accountants involved in marketing.

Early on, Mr. Kanatsu states, "This book is concerned primarily with accounting's role in the deployment of TQC in marketing and with TQC's function in accounting operations upon which business management is based" (p. 30). Near the middle he observes that, "Business is based on accounting—and since accounting is also the foundation of marketing, accounting's top level managers also participate directly in the development of company policies and plans" (p. 104). His concluding recommendation is that "while accounting departments should study the fundamentals of TQC carefully, it is even more important that they be well grounded in the basics of marketing" (p. 218).

The author advocates a strong accounting role in marketing. He believes that, "A big job of accounting's top level managers is to use statistical QC methods to analyze last year's sales and market conditions as well as to make cost breakdowns" (p. 105). What's more, "It is the accounting department's responsibility to take these trends and bits of information neglected by the sales department and teach the rest of the company what they mean and why they are important" (p. 105). Despite this "big picture" orientation the author also believes, "The accounting department has an economic and moral duty to account for everthing down to the last penny" (p. 112). What's more, "the accounting department should stress basic skills such as calculation and penmanship" (p. 121).

The book begins with the garbled statement that, "For our purposes, *accounting* is a set of activities carried out by accounting people in corporate TQC activities related to marketing" (p. 12). The definition of quality is also troubling: "Quality refers to the work standards of a product or the reliability of the work (p. 20)." Most authors operationally define quality as conformance to standards. I believe the conceptual definition should focus on meeting (exceeding?) customer's expectations.

The book includes formulas for the economic order quantity and net present value, and even introduces zero based budgeting (pp. 46–56), but contains nothing on JIT or the problem of justifying investments to achieve higher quality. While traditional cost variances are presented, the difficulties in interpreting them are not considered. While variable costing is contrasted with total costing and advocated as an important element of TQC, no mention is made of activity-based costing or costing in a JIT environment.

Important accounting tasks in TQC include: preparing budget reports (p. 81), evaluating the contribution of markets and products (pp. 81–82), developing standards, and maintaining a current procedures manuals (pp. 87–89). In a chapter on statistical methodology, attention is given to the "stratification" of data by geography, customer, and product (pp. 171–76). The discussion is typical of that in accounting texts, except that it is followed by illustrations of sales data displayed using various graphs (pp. 177–201). The chapter concludes with an analysis of total costs using scatter diagrams and equations for simple linear regression (pp. 202–04). Surprisingly, the chapter does not illustrate control charts.

Of interest are the insights into Japanese business practices woven into the text. I learned, for example, that quality circles in most Japanese companies are "a top-down system in which policy is set at the executive level" (p. 113) and that most Japanese organizations "tend to view TQC as a cost reduction process delegated to QC circles" (p. 113). Mr. Kanatsu believes quality circles are "fashioned from the ethnic identity of the Japanese people and by the familial type of management found in Japan (and that quality circles were) given life by the desire of the Japanese people to pull together to restore the nation's economy after its defeat in World War II" (p. 134).

My overall evaluation is unfavorable. While the title is appealing, the book does not measure up to this customer's expectations.

WAYNE J. MORSE
Professor of Accounting

Clarkson University

JOHN D. MARTIN and JOHN W. KENSINGER, *Exploring the Controversy Over Corporate Restructuring* (Morristown, New Jersey: Financial Executives Research Foundation, 1990, pp. vii, 124, $8.00, paper).

The 1980s were an era of unprecedented restructuring activities for corporations. For instance, there were splitups, spinoffs, splitoffs, and selloffs, as well as mergers, acquisitions, hostile takeovers, leveraged management buyouts, and so forth. As this type of activity increased, opponents (and proponents) jumped on the bandwagon to refute (and support) the restructuring of corporations.

This book is an excellent summary and review of the questions raised by the critics of restructuring. It provides answers to those critics and summarizes the results of research studies in the area.

For instance, individuals have criticized corporate restructuring for sacrificing the long-run interests of shareholders for short-run improvements. However, as the authors illustrate with their review of event studies, it has been shown that common stock prices do rise as a result of restructuring. This has been interpreted as a positive result for the shareholders by advocates of restructuring, but others argue that stock price movements may not accurately reflect the underlying value of the company. Furthermore, even if they did reflect this value, shareholders' interests may be enhanced while other groups, such as creditors and employees may be made to suffer.

The above criticisms, and others, and responses to those criticisms are discussed in a well-developed manner. This book makes a clear contribution as a reference. It could be used to present restructuring issues to individuals with limited background in this area or for those lacking training in research methods. The technical aspects of the events study methodology and the stakeholder versus the shareholder argument are left to the appendices, for those with the necessary background and interest.

One useful aspect of the book is contained in the final appendix—a discussion of on-line data services. Various data bases are available that specialize in corporate restructuring. The sources and the coverage of each of these data services are detailed. Following this appendix is a bibliography organized by the various topics discussed in the book, making it easy for the reader to obtain additional references in a given area. This book would lend itself well to classroom use in either the finance or accounting area.

JANET I. KIMBRELL
Associate Professor of Accounting

Oklahoma State University

BRIAN H. MASKEL, *Performance Measurement for World Class Manufacturing: A Model for American Companies* (Cambridge, MA: Productivity Press, Inc., 1991, pp. xxi, 408, $45).

This book provides some well-illustrated examples of performance measures alleged as useful to the emerging style of management. It is intended to provide busy managers some practical guidelines in developing consistent, useful, and relevant measures of performance as they strive for a "world class status." The contents represent a compilation from the author's consulting practice and from his experience as a Vice President of an advanced software firm.

The first three chapters recap recent management accounting themes: what is needed to be a world class organization, what is required for a performance measurement to be useful, and what is wrong with the traditional accounting model. Six chapters are devoted to a variety of measures. The illustrated methods emphasize simplicity, understandability, and continuous improvement with the key message being that performance measures must be tied to specific management objectives. The selected measures are actually used by one or more organizations.

Each chapter is started by discussing a set of contemporary management goals such as receipt of materials on time, elimination of non-value added activities, having production flexibility, achieving quality, or having good employee morale and teamwork. Then several possible reports (some being simple-to-read graphs and some being boards maintained at the work site) are discussed that will help monitor the achievement of the goal. A rough count of the number of "reports" discussed in the book is 96.

The author, in most cases, discusses the data collection system needed for the report and some of the problems that might be encountered. For example, in discussing the index, "Percentage of Standard or Common Parts," the problem of defining a common part is explored (p. 179). For several reports there is the suggestion that the data be acquired by sampling rather than by an exhaustive collection of the effect of all transactions. Also, for some reports, there is advice on how to use them (or not use them). For example, it is pointed out in the discussion of several reports to anticipate the results going in the "wrong direction" during the initial stages of a change in the way of operating. As a more specific example, Maskel points out the "average number of tasks *certified* to perform," as a measure of the degree of cross-training, will be meaningless if employees do not actually perform many of the different tasks for which they are certified (p. 200). If we use "average number of tasks performed" (instead of certified) as a measure, then we have a data collection problem. As a last example, we are warned that if we compute the "percentage of work force activity involved in quality units," then we must be careful to not use the data in a way that coerces people to participate unwillingly (p. 292). The caution sections of the book could have been improved with a better consideration of whether the reports might be artificially managed to look better than reality. (Such a discussion was plentiful when the author addressed the alleged problems of traditional management accounting. It is doubtful if all of the reports illustrated in the book are free of artificial manipulations. This may not be a concern depending on how a report is used but that is also true of traditional management accounting performance measures.)

Although one chapter considers financial performance measures, they are, in large part, shunned by Maskel. He feels they are too hard to understand. Contributing to the alleged confusion is the effect of price changes on data that are tracked over time and, for several reports, the need to allocate costs in order to compute the result.

A chapter is also devoted to the establishment of performance targets. This includes the pros and cons of having targets and the ideas of competitive benchmarking, price targeting, and the half-life of continuous improvement. Another chapter covers the production of the reports (including the need to abandon previous measures "cold turkey" and an introduction to fourth generation software report writers). A final chapter is devoted to alternative cost/management accounting techniques. These include a shift of emphasis from standards to actual and the ABC system.

The book is a welcome addition to the performance measurement literature that, to date, has been mostly general in nature. If the book succeeds in its mission, there will be radical changes in the way cost/management accounting is taught and practiced. Although intended for the practitioner the book would be a useful supplement in either an upper-level undergraduate or a graduate course.

RONALD V. HARTLEY
Professor of Accounting

Bowling Green State University

BANKS McDOWELL, *Ethical Conduct and the Professional's Dilemma: Choosing Between Service and Success* (New York: Quorum Books, 1991, pp. xi, 189, $37.95).

Professionals who are in such fields as law, medicine, nursing, accounting, securities, and insurance and who practice in a competitive, materialistic environment may be tempted to provide unnecessary services to clients or pad fees by providing excessive services along with those requested by the client. This book, written by a Professor of Law, explores this ethical dilemma and provides possible solutions.

Certain characteristics separate professions from trade occupations and provide an opportunity for this ethical dilemma to occur. These characteristics include expertise, professional education, professional organization, state licensing, and codes of ethics. The author identifies three other characteristics: (1) relationship of trust with the client, (2) internal commitment to competence and serving others, and (3) freedom from governmental regulation. Clients seek technical advice from professionals and because professionals comply with a code of ethics, society permits them to practice without strict monitoring.

Various aspects of ethics codes are examined. A code should give guidelines for resolving conflicts among the many roles and relationships of a professional. Potential role conflicts include conflicts with the client, employer, fellow professionals, members of related, competing, or subordinate professions, the profession as a group, and society. According to the author, the fewer role conflicts experienced by members of a profession, the greater the prestige that profession possesses. After analyzing the codes of several professions, including the Professional Code of Ethics (American Institute of CPAs, 1986), the author concludes that public accountants are less concerned with providing unnecessary services and most concerned with objectivity, honesty, and the reliability of their reports. Accountants wish to avoid untruthful statements and conflicts of interest that could cause ethical violations.

The remainder of the book concentrates on how to resolve the original ethical dilemma. A continuum is established containing four zones of professional practice: unnecessary services, services that are arguably unnecessary, services about which there is reasonable disagreement regarding necessity, and those services that are clearly necessary. The middle two zones cause most problems and require resolution. Revisions to codes of ethics providing better guidance and elaboration of ethical obligations could assist professionals in making appropriate judgments. Alternatively, the legal profession in the United Kingdom separates the role of advising clients (solicitor) from that of providing professional services (barrister). Other solutions include market forces eliminating marginal professionals; transference of risk to the professional through contingent fees; and increased use of peer review.

The book concludes with advice for educators noting that professional faculty are really multiprofessionals. Role conflicts exist between the role as teacher, researcher, and practitioner. Professional teachers have an ethical responsibility to prepare students to exercise autonomy in a socially responsible manner.

This book is well-written and provides good descriptions of professional characteristics and responsibilities. It prompts professionals to reflect upon their professional practice and personal lives. However, there are several weaknesses. It would appear that the author is not familiar with the role and responsibility of public accountants. The major ethical dilemma discussed, providing unnecessary services, is not the most important one faced by accounting professionals. Secondly, the cited Professional Code of Ethics (AICPA) is no longer in use, outdating the book prior to publication. If the author had consulted the original code, rather than a secondary source, the author might have been aware of the revised Code. Referencing the revised Code could have resulted in a better understanding of the public accountant's responsibility to the public and of more relevant ethical dilemmas. Additionally, this broader public responsibility and the role conflicts of the accounting profession are not perceived as casting less prestige on accountants, but rather, provide a complex, dynamic environment with responsibilities that are taken seriously and with pride.

HEIDI HYLTON MEIER
Assistant Professor of Accounting

Cleveland State University

BENNIE H. NUNNALLY, JR., D. ANTHONY PLATH, and HELENE W. JOHNS, *Corporate Lease Analysis: A Guide to Concepts and Evaluation* (New York: Quorum Books, 1991, pp. ix , 242, $49.95).

This book deals with corporate leasing, mainly from the point of view of the lessee. According to the authors, the main objectives of the book are: (1) to provide a clear delineation of the various types of leases and their advantages, disadvantages, and appropriate usage; (2) provide a thorough assessment of leasing decisions, in view of the 1986 Tax Reform Act; and (3) to accomplish the first two objectives by using an expositional style appropriate to practitioners.

The book has been well-organized into five sections. The first section, Leasing Concepts, introduces leases to the reader and provides a list of questions for which the book tries to provide answers. The second section, The Analytical Foundations of Leasing, deals with advantages and disadvantages of leasing from the lessee's standpoint. It discusses various types of leases plus gives a brief discussion of the tax implications of the lease. This section also introduces compound interest, and briefly discusses annuities, perpetuities, and effective rates of return before moving to capital budgeting and discussions of NPV, IRR, and payback.

The third section, Reporting for Leasing, discusses the tax aspects of leasing in more detail. Here the reader is given a theoretical discussion of taxes for leasing purposes, from both the point of view of the lessee and the lessor. This section also provides a good insight as to how leases should be classified and accounted for under FASB Statement No. 13.

Section Four, Evaluating Lease Transactions, gives a detailed analysis of the valuation of leasing contracts from the lessee's point of view. It examines two important finance decisions regarding the lease: capital budgeting decisions, i.e., how to select the best project (whether to lease or use an alternative), and the best wav to finance the project selected. The last section, The Role of Leasing in Modern Financial Management, discusses the contemporary theoretical and empirical research regarding lease valuation. It also examines the right and wrong reasons for leasing, some misconceptions about leasing, and also talks about some unresolved issues in leasing. The section ends with a bibliography of the leasing literature.

The book has been written in a style that makes reading simple and provides easy understanding while still giving the reader a fairly good foundation on lease financing. However, the methods are not elaborately explained with examples. There are only a few numerical examples and no problem material for the reader to work with. The examples that are provided are straightforward, in the sense that the reader does not gain a deep insight into practical problems. The book does a good job of providing separate coverage to the three main areas relevant to the leasing decision (i.e., accounting, tax, and finance). Unfortunately, no attempt is made to integrate these areas for the reader who is attempting to make a decision in the real world.

The objective achieved by the book is to provide the reader with an overview of lease concepts and analysis, covering the major aspects. But it does not provide in-depth knowledge or elaborate tools for the reader, which would help him or her to make decisions. It lacks integration, numerical examples, case studies, and problems. These would help the reader to sharpen his or her problem-solving abilities and put the knowledge to practical use. Nonetheless, some readers may find the book an excellent introductory book about leasing.

DAN PALMON
Associate Professor

Rutgers University

BOOKS RECEIVED AUGUST 11, 1991 TO NOVEMBER 5, 1991

PETER BAIDA, *Poor Richard's Legacy: American Business Values from Benjamin Franklin to Michael Miliken* (New York: William Morrow and Company, Inc., 1991, pp. 360, $11.00, paper).

AHMED BELKAOUI, *Handbook of Cost Accounting Theory and Techniques* (New York: Quorum Books, 1991, pp. xvi, 381, $75.00).

PETER W. BERNSTEIN, Editor, *The Ernst & Young Tax-Saving Strategies Guide 1991–1992* (New York: John Wiley & Sons, Inc., 1991, pp. vii, 244, $6.95, paper).

J. E. BORITZ, *The "Going Concern" Assumption: Accounting and Auditing Implications* (Toronto: The Canadian Institute of Chartered Accountants, 1991, pp. xxvii, 231, C$34.50, paper).

ANDREW H. CHEN and JOHN W. KENSINGER, *Innovations in Equity Financing* (New York: Quorum Books, 1991, pp. xv, 191, $39.95).

TOM COPELAND, TIM KOLLER, and JACK MURRIN, *Valuation: Measuring and Managing the Value of Companies* (New York: John Wiley & Sons, 1991, pp. xviii, 428, $34.95, paper).

GRAEME W. DEAN, FRANK L. CLARKE, and O. FINLEY GRAVES, *Replacement Costs and Accounting Reform in Post-World War I Germany* (New York: Garland Publishing, Inc., 1990, pp. xvii, 188, $45.00).

NICHOLAS DOPUCH and RONALD R. KING, *Experimental Tests of Auditing as a Credibility Generating Mechanism* (Vancouver: The Canadian Certified General Accountants' Research Foundation, 1991, pp. xii, 87, C$15.00 paper).

WILLIAM F. HIXSON, *A Matter of Interest: Reexamining Money, Debt, and Real Economic Growth* (New York: Praeger Publishers, 1991, pp. xxix, 274, $49.95).

M. R. LAMBERT, D. H. CARTWRIGHT, J. S. CUNNINGHAM, M. H. O'CONNOR, and M. WALSH, *Interim Financial Reporting: A Continuous Process* (Toronto: The Canadian Institute of Chartered Accountants, 1991, pp. xvii, 190, C$34.50, paper).

RICHARD MATTESSICH, *Accounting Research in the 1980s and its Future Relevance* (Vancouver: The Canadian Certified General Accountants' Research Foundation, 1991, pp. xi, 276, $22.00, paper).

BANKS MCDOWELL, *Ethical Conduct and the Professional's Dilemma: Choosing Between Service and Success* (New York: Quorum Books, 1991, pp. xi, 189, $37.95).

JANICE MONTI-BELKAOUI and AHMED BELKAOUI, *Accounting in the Dual Economy* (New York: Quorum Books, 1991, pp. xii, 159, $45.00).

ELLEN L. PAVLIK and AHMED BELKAOUI, *Determinants of Executive Compensation: Corporate Ownership, Performance, Size and Diversification* (New York: Quorum Books, 1991, pp. xii, 163, $49.95).

JOSHUA RONEN and JOSHUA LIVNAT, Editors, *Accounting and Financial Globalization* (New York: Quorum Books, 1991, pp. xiv, 167, $45.00).

ROBERT L. SHOOK, *Hardball Selling: How to Turn the Pressure On Without Turning Your Customers Off* (New York: Quill, William Morrow, 1991, pp. 224, $10.00, paper).

GEORGE SORDS, *Underwriting Democracy* (New York: The Free Press, 1991, pp. xvi, 258, $22.95).

RAJENDRA P. SRIVASTAVA, Editor, *Auditing Symposium X: Proceedings of the 1990 Deloitte & Touche/University of Kansas Symposium on Auditing Problems* (Lawrence, KS: University of Kansas, 1990, pp. ix, 186, $15.00, paper).

FREDERIC M. STINER JR., M. SUSAN STINER, and ICHIRO SHIINA, *The Accounting Profession in America* (Chiba, Japan: Chuo Gakuin University, 1991, pp. 199, Y1300, paper).

EDITORIAL POLICY AND STYLE INFORMATION

EDITORIAL POLICY

According to the policies set by the Publications Committee (which were endorsed by the Executive Committee and were published in the *Accounting Education News*, June 1987), *The Accounting Review* "should be viewed as the premier journal for publishing articles reporting the results of accounting research and explaining and illustrating related research methodology. The scope of acceptable articles should embrace any research methodology and any accounting-related subject, as long as the articles meet the standards established for publication in the journal.... No special sections should be necessary. The primary, but not exclusive, audience should be—as it is now—academicians, graduate students, and others interested in accounting research."

The primary criterion for publication in *The Accounting Review* is the significance of the contribution an article makes to the literature.

The efficiency and effectiveness of the editorial review process is critically dependent upon the actions of both authors submitting papers and the reviewers. Authors accept the responsibility of preparing research papers at a level suitable for evaluation by independent reviewers. Such preparation, therefore, should include subjecting the manuscript to critique by colleagues and others and revising it accordingly prior to submission. The review process is not to be used as a means of obtaining feedback at early stages of developing the research.

Reviewers and associate editors are responsible for providing critically constructive and prompt evaluations of submitted research papers based on the significance of their contribution and on the rigor of analysis and presentation. Associate editors also make editorial recommendations to the editor.

MANUSCRIPT PREPARATION AND STYLE

The Accounting Review's manuscript preparation guidelines follow (with a slight modification) the B-format of the *Chicago Manual of Style* (13th ed.; University of Chicago Press). Another helpful guide to usage and style is *The Elements of Style*, by William Strunk, Jr., and E. B. White (Macmillan). Spelling follows *Webster's International Dictionary*.

FORMAT

1. All manuscripts should be typed on one side of 8½ × 11" good quality paper and be double spaced, except for indented quotations.
2. Manuscripts should be as concise as the subject and research method permit, generally not to exceed 7,000 words.
3. Margins of at least one inch from top, bottom, and sides should facilitate editing and duplication.
4. To assure anonymous review, authors should not identify themselves directly or indirectly in their papers. Single authors should not use the editorial "we."
5. A cover page should show the title of the paper, the author's name, title, and affiliation, any acknowledgements, and a footnote indicating whether the author would be willing to share the data (see last paragraph in this statement).

Pagination: All pages, including tables, appendices, and references, should be serially numbered. The first section of the paper should be untitled and unnumbered. Major sections may be numbered in roman numerals. Subsections should not be numbered.

Numbers: Spell out numbers from one to ten, except when used in tables and lists, and when used with mathematical, statistical, scientific, or technical units and quantities, such as distances, weights, and measures. For example: *three days*; *3 kilometers*; *30 years*. All other numbers are expressed numerically. Generally when using approximate terms spell out the number, for example, *approximately thirty years*.

Percentages and Decimal Fractions: In nontechnical copy use the word *percent* in the text; in technical copy the symbol % is used. (See the *Chicago Manual* for discussion of these usages.)

Hyphens: Use a hyphen to join unit modifiers or to clarify usage. For example: a *well-presented analysis*; *re-form*. See *Webster's* for correct usage.

Key Words: The synopsis is to be followed by four *key words* that will assist in indexing the paper.

SYNOPSIS

A synopsis of about 400 words should be presented on a separate page immediately preceding the text. The synopsis also serves as the "introduction." The synopsis should be nonmathematical and include a readable summary of the research question, method, and the significance of the findings and contribution. The title, but not the author's name or other identification designations, should appear on the synopsis page.

TABLES AND FIGURES

The author should note the following general requirements:

1. Each table and figure (graphic) should appear on a separate page and should be placed at the end of the text. Each should bear an arabic number and a complete title indicating the exact contents of the table or figure.
2. A reference to each graphic should be made in the text.
3. The author should indicate by marginal notation where each graphic should be inserted in the text.
4. Graphics should be reasonably interpreted without reference to the text.
5. Source lines and notes should be included as necessary.

Equations: Equations should be numbered in parentheses flush with the right-hand margin.

DOCUMENTATION

Citations: Work cited should use the "author-date system" keyed to a list of works in the reference list (see below). Authors should make an effort to include the relevant page numbers in the cited works.

1. In the text, works are cited as follows: authors' last name and date, without comma, in parentheses: for example, (Jones 1987); with two authors: (Jones and Freeman 1973); with more than two: (Jones et al. 1985); with more than one source cited together (Jones 1987; Freeman 1986); with two or more works by one author: (Jones 1985, 1987).
2. Unless confusion would result, do not use "p." or "pp." before page numbers: for example, (Jones 1987, 115).
3. When the reference list contains more than one work of an author, the suffix *a*, *b*, etc. follows the date in the text citation: for example, (Jones 1987a) or (Jones 1987a; Freeman 1985b).
4. If author's name is mentioned in the text, it need not be repeated in the citation; for example, "Jones (1987, 115) says. . . ."
5. Citations to institutional works should use acronyms or short titles where practicable; for example, (AAA ASOBAT 1966); (AICPA Cohen Commission Report 1977). Where brief, the full title of an institutional work might be shown in a citation: for example, (ICAEW *The Corporate Report* 1975).
6. If the manuscript refers to statutes, legal treatises, or court cases, citations acceptable in law reviews should be used.

Reference List: Every manuscript must include a list of references containing only those works cited. Each entry should contain all data necessary for unambiguous identification. With the author-date system, use the following format recommended by the *Chicago Manual*:

1. Arrange citations in alphabetical order according to surname of the first author or the name of the institution responsible for the citation.
2. Use author's initials instead of proper names.
3. Dates of publication should be placed immediately after author's name.
4. Titles of journals should not be abbreviated.
5. Multiple works by the same author(s) should be listed in chronological order of publication. Two or more works by the same author(s) in the same year are distinguished by letters after the date.
6. Inclusive page numbers are treated as recommended in *Chicago Manual* section 8.67.

Sample entries are as follows:

American Accounting Association, Committee on Concepts and Standards for External Financial Reports. 1977. *Statement on Accounting Theory and Theory Acceptance*. Sarasota, FL: AAA.

Fabozzi, F., and I. Pollack, eds. 1987. *The Handbook of Fixed Income Securities*. 2d ed. Homewood, IL: Dow Jones-Irwin.

Demski, J. S., and D. E. M. Sappington. 1989. Hierarchical structure and responsibility accounting. *Journal of Accounting Research* 27 (Spring): 40–58.

Dye, R., B. Balachandran, and R. Magee. 1989. Contingent fees for audit firms. Working paper, Northwestern University, Evanston, IL.

Kahneman, D., P. Slovic, and A. Tversky, eds. 1982. *Judgment Under Uncertainty: Heuristics and Biases*. Cambridge, United Kingdom: Cambridge University Press.

Porcano, T. M. 1984a. Distributive justice and tax policy. *The Accounting Review* 59 (October): 619–36.

———. 1984b. The perceived effects of tax policy on corporate investment intentions. *The Journal of the American Taxation Association* 6 (Fall): 7–19.

Shaw, W. H. 1985. Empirical evidence on the market impact of the safe harbor leasing law. Ph.D. dissertation, University of Texas, Austin.

Sherman, T. M., ed. 1984. *Conceptual Framework for Financial Accounting*. Cambridge, MA: Harvard Business School.

Footnotes: Footnotes are not used for documentation. Textual footnotes should be used only for extensions and useful excursions of information that if included in the body of the text might disrupt its continuity. Footnotes should be double spaced and numbered consecutively throughout the manuscript with superscript Arabic numerals. Footnotes are placed at the end of the text.

SUBMISSION OF MANUSCRIPTS

Authors should note the following guidelines for submitting manuscripts:

1. Manuscripts currently under consideration by another journal or other publisher should not be submitted. The author must state that the work is not submitted or published elsewhere.
2. In the case of manuscripts reporting on field surveys or experiments, four copies of the instrument (questionnaire, case, interview plan, or the like) should be submitted.
3. Four copies should be submitted together with a check in U.S. funds for $50.00 for members or $100.00 for nonmembers of the AAA made payable to the American Accounting Association. Effective January 1990, the submission fee is nonrefundable.
4. The author should retain a copy of the paper.
5. Revisions must be submitted within 12 months from request, otherwise they will be considered new submissions.

COMMENTS

Comments on articles previously published in *The Accounting Review* will be reviewed (anonymously) by two reviewers in sequence. The first reviewer will be the author of the original article being subjected to critique. If substance permits, a suitably revised comment will be sent to a second reviewer to determine its publishability in *The Accounting Review*. If a comment is accepted for publication, the original author will be invited to reply. All other editorial requirements, as enumerated above, also apply to proposed comments.

POLICY ON REPRODUCTION

An objective of *The Accounting Review* is to promote the wide dissemination of the results of systematic scholarly inquiries into the broad field of accounting.

Permission is hereby granted to reproduce any of the contents of the *Review* for use in courses of instruction, as long as the source and American Accounting Association copyright are indicated in any such reproductions.

Written application must be made to the Editor for permission to reproduce any of the contents of the *Review* for use in other than courses of instruction—e.g., inclusion in books of readings or in any other publications intended for general distribution. In consideration for the grant of permission by the *Review* in such instances, the applicant must notify the author(s) in writing of the intended use to be made of each reproduction. Normally, the *Review* will not assess a charge for the waiver of copyright.

Except where otherwise noted in articles, the copyright interest has been transferred to the American Accounting Association. Where the author(s) has (have) not transferred the copyright to the Association, applicants must seek permission to reproduce (for all purposes) directly from the author(s).

POLICY ON DATA AVAILABILITY

The following policy has been adopted by the Executive Committee in its April 1989 meeting.

"An objective of (*The Accounting Review, Accounting Horizons, Issues in Accounting Education*) is to provide the widest possible dissemination of knowledge based on systematic scholarly inquiries into accounting as a field of professional, research, and educational activity. As part of this process, authors are encouraged to make their data available for use by others in extending or replicating results reported in their articles. Authors of articles which report data dependent results should footnote the status of data availability and, when pertinent, this should be accompanied by information on how the data may be obtained."

APRIL 1992 PLACEMENT ADS

The deadline for information to be included in the "Placement Information" section is two months prior to the date of publication. The REVIEW *is published on the first day of January, April, July, and October. Placement advertising and replies should be sent to the American Accounting Association, 5717 Bessie Drive, Sarasota, FL 34233. ATTENTION: ADVERTISING DEPT. All replies to box numbers should be on separate sheets to facilitate forwarding. There is no charge for this service to AAA members and associate members seeking employment. Others will be charged a fee of $150 per insertion. ADS SHOULD NOT EXCEED 150 WORDS.*

#201 THE UNIVERSITY OF CALGARY Faculty of Management invites applications for positions at all levels in the areas of taxation, auditing, financial and managerial accounting. Candidates should have a Ph.D., or D.B.A. with a major in Accounting. ABDs in advanced dissertation stage will be considered. Research and publication requirements will be commensurate with rank upon appointment. Salary is competitive and computer assistance and research support are available. Starting date will be determined by mutual agreement. The Faculty of Management offers undergraduate MBA and Ph.D. programs in Accounting. The University of Calgary has an employment Equity Program and encourages applications from all qualified candidates, including women, aboriginal people, visible minorities and people with disabilities. Interested individuals are asked to mail applications to: Rashmi Thakkar, Ph.D., ACA, CMA, LLB, Chair, Accounting, Faculty of Management. The University of Calgary, 2500 University Drive N.W., Calgary, Alberta T2N 1N4, Canada.

#202 THE UNIVERSITY OF MARYLAND European and Asian Divisions seek excellent teachers for undergraduate courses on U.S. military bases overseas. Minimum teaching experience, and U.S. citizenship. Renewable annual appointment to begin August 1992. Benefits include transportation and military base privileges. Frequent travel and the cost of schooling make these positions difficult for those with children. Send resumé to: Dr. Ralph E. Millis, The University of Maryland University College, College Park, MD 20742-1642. AA/EEO.

#203 THE UNIVERSITY OF NEW ORLEANS invites applications for anticipated positions in the areas of Auditing, Taxation, or Financial, or some combination thereof. Rank is at the Assistant Professor level. Candidates must have the Ph.D. (accounting major). Salary is negotiable and competitive. The University offers the M.S. in accounting and the M.S. in accounting (tax option), as well as the MBA. Aproximately 550 undergraduate accounting majors are enrolled. The Department has separate AACSB accounting accreditation at both the graduate and undergraduate levels. Candidates should send a resumé to: Professor Philip J. Harmelink, Chairman, Department of Accounting, College of Business Administration, University of New Orleans, New Orleans, LA 70148. UNO is a member of the Louisiana State University System and an Affirmative action, Equal Opportunity Employer.

#204 Very prolific tenured Associate Professor (Ph.D.) seeking a full-time (academic, administrative, research and/or a combination) position in an Accounting over the past 10 years (at the University of Miami) includes: Financial/Managerial/Cost/Forensic Auditing and Accounting, Accounting Research Methods/Expert System and Artificial Intelligence/Quantitive Methods/Information Systems in Accounting as well as Computer Auditing/Programming for graduate/undergraduate/executive and professional courses. I have enjoyed teaching, advising students, researching and publishing, supervising thesis, consulting and interacting with the community/professionals.

Meeting the Challenges
OF
ACCOUNTING EDUCATION
IN THE
'90s

◆

IRWIN
THE ACCOUNTING PUBLISHER

is proud to present our 1992 titles:

Management Control Systems 7/e
Robert N. Anthony
John Dearden
Vijay Govindarajan

Accounting for Governmental and Non-Profit Entities 9/e
Leon E. Hay
Earl R. Wilson

Accounting Theory 5/e
Eldon Hendriksen
Michael van Breda

Principles of Auditing 10/e
O. Ray Whittington
Kurt Pany
Walter B. Meigs
Robert F. Meigs

Accounting Principles 5/e
Roger H. Hermanson
James Don Edwards
Michael W. Maher

Financial Accounting 5/e
Roger H. Hermanson
James Don Edwards

College Accounting
Kenneth A. Koerber

Financial Accounting 5/e
Kermit D. Larson
Paul B. W. Miller

Intermediate Accounting
Revised and Standard Editions
Thomas R. Dyckman
Roland E. Dukes
Charles J. Davis

Also available from IRWIN, annual editions of:
The FASB Volumes
Original Pronouncements, Volumes I & II
Current Text, Volumes I & II
Statements of Financial Accounting Concepts

For more information write or call
your IRWIN representative or
Faculty Service at 1-800-323-4560.

IRWIN
1818 RIDGE ROAD HOMEWOOD, IL 60430

ACCOUNTING TITLES

AUDITING

NEW	**Auditing** **Fifth Edition**	Thomas D. Hubbard, *University of Nebraska—Lincoln* Johnny R. Johnson, *The American University*

FINANCIAL AND MANAGERIAL ACCOUNTING

	Principles of Accounting **Sixth Edition**	James J. Benjamin, *Texas A&M University* Arthur J. Francia, *University of Houston* Robert H. Strawser, *Texas A&M University*
	Financial Accounting **Tenth Edition**	James J. Benjamin, *Texas A&M University* Arthur J. Francia, *University of Houston* Robert H. Strawser, *Texas A&M University*
	Financial Accounting **Revised Edition**	Vincent C. Brenner, *Louisiana State University* Ronald M. Copeland, *Northeastern University* Paul E. Dascher, *Drexel University*
NEW	**Financial Accounting Text and Cases**	Paul R. Berney, *Yale School of Organization and Management* Jimie Kusel, *University of Arkansas—Little Rock*
	Managerial Accounting **Eighth Edition**	Arthur J. Francia, *University of Houston* Mattie C. Porter, *University of Houston—Clear Lake* Robert H. Strawser, *Texas A&M University*
	Managerial Accounting **Third Edition**	Ronald M. Copeland, *Northeastern University* Paul E. Dascher, *Drexel University* Jerry R. Strawser, *University of Houston*
	Managerial Cost Accounting: Planning and Control **Revised Edition**	William L. Ferrara, *Stetson University* Frank P. Dougherty, *Saint Mary's University* Germain B. Boer, *Vanderbilt University*
NEW	**Cost Accounting for Managerial Planning, Decision Making and Control** **Fourth Edition**	Woody M. Liao, *University of California—Riverside* James L. Boockholdt, *Samford University*
	The Executive's Guide to Management Accounting and Control Systems **Fourth Edition**	**For MBAs and Professional Development** William Rotch, *University of Virginia* Brandt R. Allen, *University of Virginia* C. Ray Smith, *University of Virginia*
NEW	**Income Tax for Accounting Principles** **1992 Edition**	Benny R. Copeland, *University of North Texas* John E. Price, *University of North Texas*
NEW	**Practice Case in Financial Statement Preparation and Analysis** **The Keller Corporation**	Francis E. Kearns, *Rochester Institute of Technology* Robert P. Derstine, *Villanova University* Ronald J. Heufner, *State University of New York—Buffalo*
NEW	**Practice Problem in Financial Accounting** **Music Machines, Incorporated** **Revised Edition**	Thomas P. Edmonds, *University of Alabama—Birmingham* Margaret L. Shelton, *University of Houston—Downtown*
NEW	**Practice Problem for Intermediate Accounting I** **Monroe & Sons, Incorporated** **Third Edition**	Thomas P. Edmonds, *University of Alabama—Birmingham* Margaret L. Shelton, *University of Houston—Downtown*

DAME
PUBLICATIONS, INC

7800 Bissonnet—Suite 415
Houston, TX 77074
(713) 995-1000 FAX (713) 995-9637

When requesting examination copies, please indicate course title and text(s) currently in use.